

Kraftmikroskopie als Werkzeug zur Charakterisierung elektrischer Oberflächeneigenschaften

Von der Fakultät für Mathematik, Naturwissenschaften und Informatik

der Brandenburgischen Technischen Universität Cottbus

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

(Dr. rer. nat.)

genehmigte Dissertation

vorgelegt von

Diplom Physiker

Markus Ratzke

geboren am 11.01.1976 in Finsterwalde (Deutschland)

Gutachter: Prof. Dr. rer. nat. Jürgen Reif

Gutachter: Prof. Dr. rer. nat. Götz Seibold

Gutachter: Dr. rer. nat. Mario Birkholz

Tag der mündlichen Prüfung: 07.11.2011

Zusammenfassung

Die stetig fortschreitende Miniaturisierung in der Halbleiterindustrie macht es notwendig, Oberflächenparameter mit Auflösung im Nanometerbereich zu messen und auch abzubilden. Von größtem Interesse sind hierbei das Oberflächenpotential und die Kapazität der oberflächennahen Bereiche, da diese Aussagen über die elektronische Struktur erlauben.

Hierbei muss großes Augenmerk auf die Möglichkeit der zerstörungsfreien und präparationsarmen Messung gelegt werden, da jede Behandlung der zu untersuchenden Materialien deren Oberflächeneigenschaften ändert.

Im Rahmen dieser Arbeit wurden auf der Atomkraftmikroskopie basierende Methoden sowohl experimentell als auch mit Hilfe von Simulationen auf ihre Anwendbarkeit für die Untersuchung von Halbleiteroberflächen evaluiert.

Es stellt sich heraus, dass die kontaktfreien Methoden „Scanning Kelvin Probe Microscopy“ und „Scanning Capacitance Microscopy“ sehr gut geeignet sind, um die elektronische Struktur der Probenoberfläche qualitativ zu beurteilen. Allerdings muss für quantitative Aussagen ein recht großer rechentechnischer Aufwand betrieben werden.

Abstract

The still ongoing trend of miniaturization in the semiconductor industry creates the demand for experimental methods allowing a mapping of certain properties of surfaces with resolution in the nanometre range. Matters of particular interest are the surface potential and the capacitance of near surface regions, since these parameters give access to electronic properties.

However, of major interest are non-destructive experimental methods, especially these with minor requirements on sample preparation allowing investigations of pristine surfaces.

In this study, atomic force microscopy-based methods are evaluated both experimentally and with the help of simulations for their applicability to the investigation of semiconductor surfaces.

It turns out that the non-contact methods, "Scanning Kelvin Probe Microscopy" and "Scanning Capacitance Microscopy" are very well suited to estimate the electronic structure of the sample surface qualitatively. However, for a quantitative assessment a fairly large effort of computing has to be carried out.

Schlagwörter:

Atomkraftmikroskopie, Silizium, Scanning Kelvin Probe Microscopy, Scanning Capacitance Microscopy

Keywords:

Atomic Force Microscopy, Silicon, Scanning Kelvin Probe Microscopy, Scanning Capacitance Microscopy

Inhaltsverzeichnis

Zusammenfassung.....	2
Vorwort	5
1. Motivation	7
2. Konzeptionelle Grundlagen	9
2.1 Rastersondenmikroskopie	9
2.2 Das experimentelle Setup	17
2.3 Numerische Behandlung von Randwertproblemen	23
2.4 Die Methode der finiten Elemente.....	34
2.5 Numerisches Vorgehen	37
2.6 Numerische Fehler	41
3. Grundlegende Wechselwirkungen im System Spitze-Probe	45
3.1 Generelle Abhängigkeiten im System Spitze-Metall	45
3.2 Einfluss von dielektrischen Schichten auf die Messwerte.....	52
3.3 Untersuchungen zur Auflösung in einfachen Systemen	56
3.4 Komplexe Systeme – Metallische Spitze / Halbleiteroberfläche.....	66
3.5 Einfluss von Grenzflächenladungen auf die Messwerte	74
3.6 Die Auflösung im Fall halbleitender Proben.....	77
3.7 Methodischer Vergleich zwischen Kapazitäts- und Kraftmessung.....	81
4. Die Transferfunktion.....	85
4.1 Näherungsweise analytische Bestimmung der Transferfunktion	85
4.2 Numerische Bestimmung der Transferfunktion	96
5. Ausgewählte experimentelle Ergebnisse	113
5.1 Allgemeines zu den Methoden – Typische Kontraste	113
5.2 Gebondete Silizium-Wafer	115
5.3 Dotiergitter in Silizium.....	123
Zusammenfassung und Ausblick	135
Danksagung	137
A Grundlegende Gleichungen für die elektrostatische Beschreibung von Silizium	139
A.1 Die Poissongleichung.....	139
A.2 Häufig benutzte Näherungen und typische Längenskalen	140
A.3 Folgerungen aus der Lösung – Die Kapazität der Oberfläche	143
A.4 Der Readzylinder	145
B Intermolekulare Kräfte im System Probe-Sonde.....	147

B.1 Anziehende Kräfte	147
B.2 Abstoßende Kräfte	151
B.3 Das komplette Modell	152
C Der belichtete Halbleiter	153
C.1 Generation und Rekombination	153
C.2 Mathematische Beschreibung des Ladungsträgerflusses	156
C.3 Das Drift-Diffusions-Modell	158
D Multipolentwicklungen	163
Literaturliste	167

Vorwort

Seit jeher ist die Menschheit daran interessiert, den Aufbau der Materie zu ergründen. Ein Grund dafür ist schlichter Forschergeist, ein anderer die Notwendigkeit, künstlich erzeugte Strukturen auf ihre Funktionsfähigkeit hin zu untersuchen. Da das menschliche Auge nur begrenzt zur Betrachtung kleiner Gegenstände geeignet ist, wurden und werden technische Hilfsmittel erdacht, um den zugänglichen Größenordnungsbereich zu erweitern.

Die Geschichte dieser Hilfsmittel, der Mikroskope (aus dem Griechischen: μικρός - klein und σκοπεῖν - betrachten), begann schon in der Römerzeit. Damals wurden wassergefüllte Glasschalen als Vergrößerungsgläser benutzt. Um 1600 wurden zusammengesetzte Mikroskope entwickelt, die bis zu 50fache Vergrößerung¹ erreichten. 1880 bildeten schließlich die Berechnungen Ernst Abbes [1] die Grundlage für den Bau optischer Geräte, deren Auflösung nicht durch verfahrenstechnische Bedingungen limitiert wurden, sondern durch die physikalischen Gesetze der Beugung. Es ergibt sich, dass das Auflösungslimit bei Verwendung von sichtbarem Licht in der Größenordnung der verwendeten Wellenlänge, also zwischen 300 und 700 nm, liegt. Verglichen mit den „interessanten“ Strukturgrößen erscheint diese Zahl allerdings inadäquat, weshalb um 1930 andere Methoden der Abbildung gesucht wurden.

Ernst Ruska gelang es 1932 [2] ein Elektronenmikroskop zu entwickeln, welches einen gebündelten Elektronenstrahl zur Darstellung von metallischen Oberflächen benutzte. Die erreichte Auflösung lag bei etwa 50 nm, also um einen Faktor 10 verbessert gegenüber der Lichtmikroskopie. Dieser Prototyp ist der Urvater der Transmissionselektronenmikroskope, welche dünne Folien der zu beobachteten Probe durchstrahlen. Die Auswertung des Streuungsmusters der Elektronen ermöglicht beispielsweise eine Darstellung der Dichteverteilung mit einer Auflösung von einigen Å, weshalb TEMs heutzutage sehr oft für die hochauflösende Abbildung der atomaren Struktur von Festkörpern benutzt werden.

Auch können die vom Elektronenstrahl erzeugten Sekundärelektronen zur Abbildung genutzt werden. Die so konzipierten Rasterelektronenmikroskope ermöglichen ebenfalls Auflösung im nm-Bereich und erfreuen sich aufgrund des einfacheren Aufbaus ebenfalls großer Beliebtheit.

Es dauerte etwa 50 Jahre, bis ein weiterer großer Sprung in der Mikroskopietechnik zu verzeichnen war: Gerd Binnig und Heinrich Rohrer entwickelten 1981 das Rastertunnelmikroskop [3], den Vorläufer aller Rastersondenmethoden. Es gelang, einzelne Atome auf Oberflächen mittels Abtastung mit einer feinen metallischen Sonde abzubilden. Da das Regelsignal hierbei der Tunnelstrom zwischen Spitze und Probe ist, muss die Probe leitend sein. Das ebenfalls von Binnig entwickelte Atomkraftmikroskop (AFM) [4] umgeht diese Limitierung durch eine direkte Messung der intermolekularen Kräfte zwischen den Atomen der Oberfläche und denen der Spitze. Da diese Kräfte unter anderem auch elektrostatischer Natur sind, kann ein AFM-System dahingehend modifiziert werden, um zur, zumindest qualitativen, Abbildung elektrischer Signale benutzt zu werden.

¹ Die Vergrößerung von Mikroskopen wird als Verhältnis der scheinbaren Größe des betrachteten Gegenstandes in der deutlichen Sichtweite (etwa 25 cm) zur tatsächlichen Größe definiert. Legt man ein Auflösungsvermögen des Auges von etwa einer Bogenminute zugrunde, ergeben sich für das erwähnte Mikroskop Strukturgrößen von ca. 50 µm.

Vorwort

Eine direkte Umsetzung makroskopischer Methoden, wie zum Beispiel Kapazitäts-Spannungsmessungen an Halbleitern, ist aufgrund der zumeist nötigen Kontakte und der damit verbundenen Zerstörung der ursprünglichen Oberflächenbeschaffenheit schwierig. Deshalb soll in dieser Arbeit die Möglichkeit untersucht werden, ob und wie gut kontaktfreie AFM-basierte Methoden dazu geeignet sind, elektrische Eigenschaften im Nanometerbereich zu untersuchen.

1. Motivation

Triebkraft für die Entwicklung vieler neuer Fertigungs- und Untersuchungsmethoden ist zunehmend die Halbleiterindustrie. In den vergangenen etwa 50 Jahren gab es in dieser Branche den stetigen Trend der Verkleinerung der Strukturen, was letztendlich höhere Ausbeute und bessere Leistungsfähigkeit der gefertigten Bauelemente bedeutet. Verfolgen kann man diesen Trend an beliebigen Kenngrößen zum Beispiel quasi alltäglicher Technologie, wie der Anzahl von Transistoren pro Prozessor. Allgemein akzeptiert ist das Mooresche Gesetz, das eine Verdopplung der Komplexität etwa alle zwei Jahre prognostiziert (siehe Abbildung 1.1).

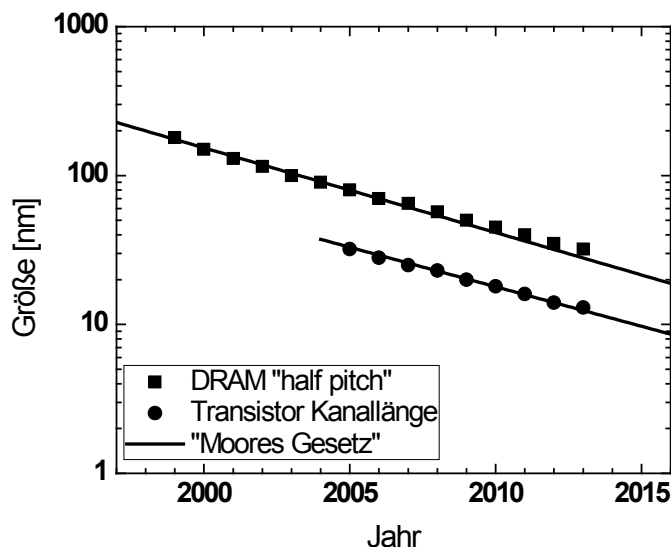


Abbildung 1.1: Typische Größenordnungen in der Halbleiterindustrie, hier der halbe Abstand zwischen Leiterbahnen in Speicherzellen (DRAM) und Kanallängen in Transistoren. Offenbar gibt es exponentielles Schrumpfen der Strukturen mit Halbierung etwa alle zwei Jahre (nach ITRS [5]).

Im Fahrwasser dieser Komplexitätssteigerung steigt natürlich auch das Interesse an bildgebenden Verfahren mit sehr großer Auflösung und Vielseitigkeit. Wie bereits im Vorwort erwähnt, erscheinen sondenbasierte Methoden durchaus vielversprechend. Eine Nutzung dieser Methoden, vor allem für die Charakterisierung halbleitender Proben aus verschiedensten Anwendungsbereichen (Photovoltaik, Mikroelektronik), setzt allerdings Folgendes voraus:

- 1.) Die Mechanismen müssen gut verstanden werden. Die Fragestellung: „Was wird wie abgebildet?“, muss zwingend geklärt werden.
- 2.) Die Sensitivität muss bekannt sein.
- 3.) Das Auflösungsvermögen muss beschrieben werden.
- 4.) Möglichkeiten zur Vermeidung von Signalverfälschungen müssen bekannt sein. Falls diese unvermeidbar sind, sollten sinnvolle Korrekturmöglichkeiten angeboten werden.

Die Klärung der genannten Problemstellungen erfordert sowohl experimentelle Auswertung von Beispielpuben als auch den Aufbau eines theoretischen Hintergrundes. Dies sind sogleich die Säulen der vorliegenden Arbeit. So wird in Kapitel 3 geklärt, welche Probeneigenschaften abgebildet werden können, Kapitel 4 beschäftigt sich mit der Frage der Auflösung. Die gewonnenen Aussagen werden in Kapitel 5 mit experimentellen Ergebnissen verglichen.

Motivation

2. Konzeptionelle Grundlagen

In diesem Kapitel sollen sowohl die verwendeten experimentellen und numerischen Methoden als auch ihre Grundlagen erläutert werden. Im experimentellen Bereich wird die Rastersondenmikroskopie, mit besonderem Augenmerk auf die tatsächlich genutzten Methoden, dargestellt. Im Abschnitt über die numerischen Grundlagen werden Herangehensweisen zur numerischen Behandlung von Randwertproblemen erläutert.

2.1 Rastersondenmikroskopie

In diesem Abschnitt sollen kurz die Entwicklung, der typische Aufbau und die Anwendungsbereiche von Rastersondenmikroskopen (Scanning Probe Microscopes, SPMs) erläutert werden. Der Ursprung aller heutigen benutzten SPMs ist im Rastertunnelmikroskop (Scanning Tunneling Microscope, STM) zu sehen, welches um 1980 im IBM-Forschungslabor Rüschlikon von Gerd Binnig und Heinrich Rohrer entwickelt wurde. Dieses Gerät nutzt die starke Abstandsabhängigkeit des quantenmechanischen Tunnelstroms², um eine atomar spitze Sonde in fester Höhe über eine Probenoberfläche zu führen. Ein angeschlossener Regelkreis verändert dazu die Auslenkung eines Piezokristalls im sub-nm Bereich (Abbildung 2.1).

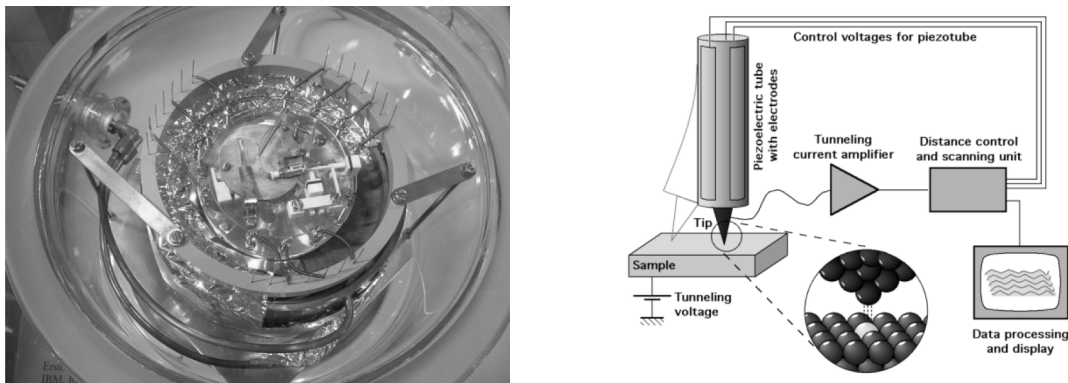


Abbildung 2.1: Typischer Aufbau eines Rastertunnelmikroskops: Das von Binnig und Rohrer gebaute erste STM (links) und eine stark vereinfachte Funktionsskizze (rechts). Quelle: [6].

Obwohl die Methode relativ einfach Auflösung im atomaren Bereich erlaubt (zusätzlich zu der erhaltenen Reliefinformation sind auch Untersuchungen der elektronischen Struktur möglich), hat sie doch einige prinzipielle Nachteile: Als Proben kommen nur leitfähige Materialien in Frage und das Experiment muss in einem guten Hochvakuum in einem möglichst schwingungsfreien Umfeld durchgeführt werden. Die hohen Anforderungen an Probe und Umfeld können behoben werden, wenn zur Abbildung zum Beispiel die Adhäsions- und Kohäsionskräfte zwischen der Probe und einer wiederum sehr feinen Spitze genutzt werden.

² Der Tunnelstrom ist proportional zur Tunnelwahrscheinlichkeit, welche für den Fall einer

dreiecksförmigen Potentialbarriere folgende Form annimmt: $\Theta = \exp\left(-\frac{4}{3} \frac{\sqrt{2qm^*}}{\hbar} \frac{d \cdot \Phi_B^{\frac{3}{2}}}{U}\right) = \exp\left(-\frac{d}{C}\right).$

Die für die Messung wichtigste Abhängigkeit ist die vom Abstand d , wobei die eingeführte Konstante C in der Größenordnung von einem Å liegt.

Konzeptionelle Grundlagen

Umgesetzt wurde diese Idee in den sogenannten Rasterkraftmikroskopen (Scanning Force Microscope, SFM; Atomic Force Microscope, AFM), welche konzeptionell ebenfalls von Gerd Binnig entwickelt wurden. Da die experimentellen Daten für diese Arbeit mit einem AFM ermittelt wurden, soll auf den Aufbau und die Funktion etwas genauer eingegangen werden.

Rasterkraftmikroskopie

Grundlegend für das Verständnis der Rasterkraftmikroskope ist das Wissen über die auftretenden Kräfte. Üblicherweise wird zur Erklärung ein Lennard-Jones-Potential benutzt, welches langreichweitige anziehende Dipol-Dipol-Wechselwirkungen (van-der-Waals-Kräfte) und kurzreichweitige abstoßende Kräfte verbindet und für zwei Teilchen folgende Form hat:

$$V(r) = \varepsilon \cdot \left(\left(\frac{r_0}{r} \right)^{12} - 2 \left(\frac{r_0}{r} \right)^6 \right) \quad (2.1.1)$$

Darin sind ε die Tiefe der entstehenden Potentialmulde, r_0 der Gleichgewichtsabstand und r der Abstand der Teilchen. Abbildung 2.2 zeigt den prinzipiellen Verlauf in normierten Variablen:

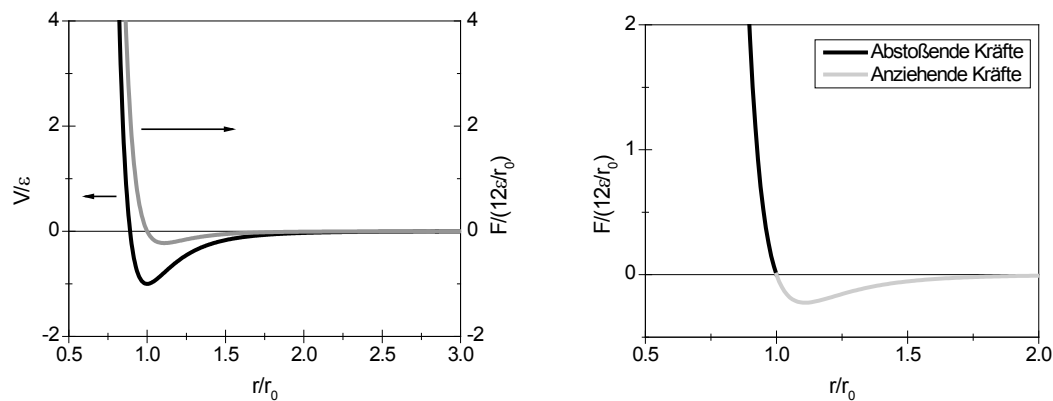


Abbildung 2.2: Verlauf des Lennard-Jones-Potentials und der resultierenden Kraft dargestellt in normierten Koordinaten (links). Verdeutlichung der verschiedenen Bereiche (rechts).

Ebenfalls dargestellt in Abbildung 2.2 ist der Verlauf der sich aus dem Potential ergebenden Kraft F :

$$F = -\frac{\partial V}{\partial r} = \frac{12\varepsilon}{r_0} \left(\left(\frac{r_0}{r} \right)^{13} - \left(\frac{r_0}{r} \right)^7 \right) \quad (2.1.2)$$

Diese Kraft zeigt für $r < r_0$ einen stark abstoßenden Charakter und für $r > r_0$ einen anziehenden. Obwohl es für die Beschreibung des Funktionsprinzips des AFM nicht notwendig ist, soll darauf hingewiesen werden, dass der Verlauf von Potential und Kraft in realen Systemen, also allen Systemen, die aus mehr als zwei Teilchen bestehen, komplizierter ist. Durch das notwendige Summieren aller auftretenden Kräfte ändern sich die Exponenten in (2.1.2), die typische Form bleibt aber erhalten. Nähere Ausführungen zu diesem Thema finden sich in Anhang B. Wie dort gezeigt, ergeben sich Kräfte von einigen nN.

Die Detektion dieser Kräfte erfolgt in allen Kraftmikroskopen durch die Auswertung der Verbiegung eines dünnen Balkens, des sogenannten Cantilevers. Dieser wird üblicherweise lithografisch aus Silizium hergestellt und hat Ausdehnungen in der Größenordnung von $100\text{ }\mu\text{m} \times 30\text{ }\mu\text{m} \times 2\text{ }\mu\text{m}$ (Länge, Breite, Dicke). Ausgehend von der Theorie der Balkenbiegung ergeben sich effektive Federkonstanten von etwa 1 N/m beziehungsweise typische Verbiegungen von etwa 1 nm . Die Erfassung dieser Verbiegung geschieht wiederum über ein System bestehend aus einem an der Oberfläche des Cantilevers reflektierten Laserstrahl und einer Quadrantendiode zur Messung relativer Beleuchtungsstärken (siehe Abbildung 2.3a).

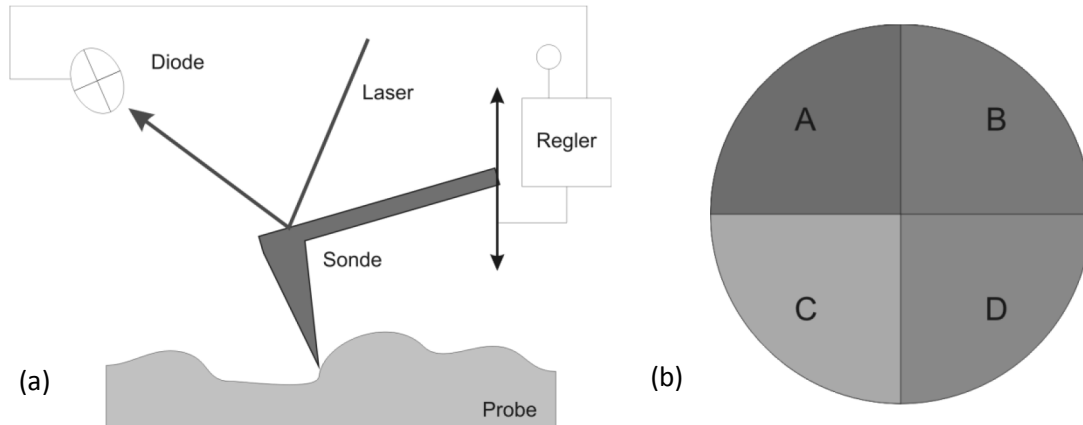


Abbildung 2.3: Typisches AFM-Setup (a) und schematische Abbildung einer Quadrantendiode (b). Der auf der Oberfläche des ausgelenkten Cantilevers reflektierte Laserstrahl erzeugt auf der Diode positionsabhängige Signale, welche zur Regelung/Detektion der Verbiegung genutzt werden.

Der wesentliche Vorteil dieser Art der Auslenkungsmessung liegt in der Verstärkung. So erzeugt eine Auslenkung w des Cantileverendes unter Annahme kleiner Winkel folgende Verschiebung Δz des Strahlmittelpunkts auf der Diode:

$$\Delta z = \frac{4}{3} \frac{D}{L} w \quad (2.1.3)$$

In (2.1.3) steht dabei L für die Länge des Cantilevers und D für den Abstand zwischen Diode und Sonde. Für übliche Werte ergeben sich so leicht Verstärkungsfaktoren in der Größenordnung von 1000, was die typischen Auslenkungen von einigen nm in den gut messbaren Bereich von μm verschiebt. Die Diode selbst wird zur Messung von Differenzsignalen konfiguriert, gemäß Abbildung 2.3b also zum Beispiel $(A+B)-(C+D)$, wobei die Buchstaben exemplarisch für die gemessenen Ströme genutzt werden. Unter der Annahme eines Gaußschen Strahlprofils lässt sich der Wert dieses Differenzsignals angeben zu:

$$(A + B) - (C + D) \sim \text{erf} \left(\frac{\Delta z}{w_{\text{Strahl}}} \right) \quad (2.1.4)$$

Diese Funktion ist linear in Δz , falls sich dieser Wert im Bereich des Strahlradius w_{Strahl} bewegt.

Angeschlossen an den Detektor ist im Weiteren ein Regelkreis, der, abhängig vom gewählten Betriebsmodus, durch Variation der Höhe des festen Cantileverendes eine Regelgröße auf einem bestimmten, festen Wert zu halten versucht. Regelgrößen können dabei die Verbiegung (Contact-Modus) oder die Amplitude der Schwingung des freien Endes (Non-Contact-Modus)

sein. Da die direkte Messung der Verbiegung aufgrund ihrer Störanfälligkeit gegen äußere Faktoren aufwendig ist, wird meist ein schwingungsbasierter Aufbau benutzt, weshalb im Folgenden kurz die Dynamik der Cantileverschwingungen beschrieben werden soll.

Die Dynamik des Spitze-Probe-Systems

Die Beschreibung der Dynamik des Systems Spitze-Probe ist von grundlegender Bedeutung sowohl für das Verständnis der Arbeitsweise eines typischen AFM-Aufbaus als auch für die der „Electrostatic Force Microscopy“ (EFM).

Aufgrund der starken Abstandsabhängigkeit aller auftretenden Kräfte, also der elektrostatischen Kräfte und der Atom-Atom-Wechselwirkungen, kann das durchaus komplizierte Konstrukt (siehe Abbildung 3.1) bestehend aus Cantilever, Spitze und Apex für die Beschreibung deutlich vereinfacht werden. Dazu werden alle Kräfte als punktförmig, wirkend am freien Ende des Cantilevers, beschrieben und die Biegung dieses Cantilevers auf die Messung der Verbiegung am freien Ende reduziert³. Diese Näherungen sind gerechtfertigt durch die Form der Differentialgleichungen, welche statische und dynamische Verbiegung eines Balkens beschreiben (die Entwicklung dieser wird Euler und Bernoulli zugeschrieben [7]):

$$\begin{aligned} (a) \quad EI w^{(4)}(x) &= q(x) \\ (b) \quad EI \frac{\partial^4 w(x,t)}{\partial x^4} + \mu \frac{\partial^2 w(x,t)}{\partial t^2} &= 0 \end{aligned} \quad (2.1.5)$$

In (2.1.5) bedeuten E den Elastizitätsmodul des verwendeten Materials, I das axiale Flächenträgheitsmoment, welches die Verbiegungssteifigkeit des Querschnittes widerspiegelt, q die Last (Kraft pro Länge), μ die Massenbelegung (Masse pro Länge) sowie w die orts- und zeitabhängige Verbiegung. Es erscheint sinnvoll, zunächst zwei statische Probleme exemplarisch zu besprechen: Die Belastung mit einer konstanten Flächenlast q und die Belastung mit einer punktförmigen Kraft F . Da die Lösung der Differentialgleichungen trivial ist und lediglich die Randbedingungen für freie Enden (verschwindende Kraft, verschwindendes Drehmoment) und feste Enden (verschwindende Verbiegung, verschwindender Ablenkwinkel) eingeprägt werden müssen, sollen hier nur die Lösungen angegeben werden:

$$\begin{aligned} (a) \quad w(L) &= \frac{1}{8} \frac{qL^4}{EI} \\ (b) \quad w(L) &= \frac{1}{3} \frac{FL^3}{EI} \end{aligned} \quad (2.1.6)$$

Offenbar sind beide Lösungen bis auf einen konstanten Faktor identisch, was die Reduktion auf den Fall der punktförmigen Kraft motiviert. Die offensichtliche Proportionalität zwischen Auslenkung w und Kraft F legt es zudem nahe, die von Federn bekannte Formulierung $F=kx$, mit einer von der Geometrie bestimmten Federkonstante k , zu benutzen.

³ Typische AFM-Systeme messen eigentlich nicht die Verbiegung am freien Ende, sondern den Verbiegungswinkel an dieser Stelle. Erfreulicherweise ist dieser aber direkt proportional zur Verbiegung, was wiederum ein zumindest sprachliches Gleichsetzen der Größen erlaubt.

Da Gleichung (2.1.5)b nach Ort und Zeit separiert werden kann und die resultierende Differentialgleichung für den zeitabhängigen Teil eine Schwingungsgleichung repräsentiert, kann die Dynamik mit folgender Gleichung modelliert werden:

$$m^* \ddot{w} + \beta \dot{w} + kw = F_{ext}(z, w, t) \quad (2.1.7)$$

Die in (2.1.7) eingeführten Konstanten m^* , β und F_{ext} stehen dabei für eine effektive Masse des Systems, eine empirisch bestimmbare Dämpfung und eine externe Kraft, welche zu Schwingungen des Cantilevers führt und im Allgemeinen von der Auslenkung w , dem Spitze-Probe-Abstand z und der Zeit t abhängt.

Der ortsabhängige Teil der separierten Version von Gleichung (2.1.5)b definiert die Form der Schwingung, die üblicherweise als Mode bezeichnet wird. In Abbildung 2.4 sind einige dieser Moden dargestellt. Unterscheidbar sind sie durch unterschiedliche Komplexität, hier zum Beispiel durch die Anzahl der Knoten und ihre Frequenz. In allen dieser Arbeit zugrunde liegenden Experimenten wurde die niedrigste Resonanzfrequenz des Systems gewählt, somit die Mode, die in Abbildung 2.4 (links) veranschaulicht ist.

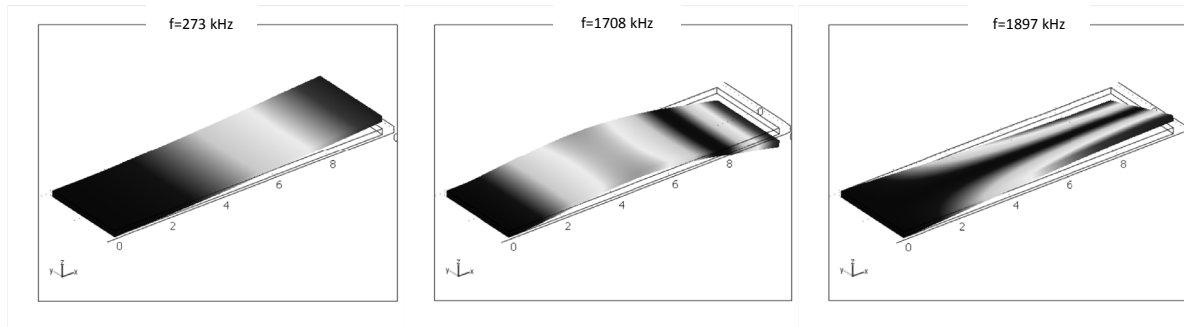


Abbildung 2.4: Verschiedene Moden eines für AFM-Systeme typischen Cantilevers. Die Grundmode (links) zeichnet sich durch lediglich einen Knoten am festen Ende und die niedrigste Frequenz aus. Im Gegensatz dazu erscheinen die höheren Moden deutlich komplexer. Die Grauskalierung versinnbildlicht in allen Fällen die Gesamtverschiebung.

Da aus den beschriebenen praktischen Gründen die geometrische Form der Moden vernachlässigt werden kann, sollen nunmehr typische Eigenschaften von Gleichung (2.1.7) besprochen werden. Zunächst wird dazu die Gleichung durch die Masse m^* dividiert und die Dämpfung durch einen Qualitätsfaktor Q beschrieben:

$$\ddot{w} + \frac{\omega_0}{Q} \dot{w} + \omega_0^2 w = \frac{F_{ext}(w, z, t)}{m^*} \quad (2.1.8)$$

Die Kreisfrequenz ω_0 ist dabei wie üblich definiert. Die Einführung einer skalierten Zeit $\tau = \omega_0 t$ ermöglicht eine etwas einfachere Darstellung, wobei die Ableitung nach τ mit „ \prime “ symbolisiert ist:

$$\begin{aligned} \omega_0^2 w'' + \frac{\omega_0^2}{Q} w' + \omega_0^2 w &= \frac{F_{ext}(w, z, \tau)}{m^*} \Leftrightarrow \\ w'' + \frac{1}{Q} w' + w &= \frac{F_{ext}(w, z, \tau)}{k} = A_{ext}(w, z, \tau) \end{aligned} \quad (2.1.9)$$

Abhängig von der generierenden Funktion A_{ext} auf der rechten Seite von (2.1.9) lassen sich verschiedene typische Situation unterscheiden:

1.) Mechanische Anregung in statischem Kraftfeld

Für diesen Fall lässt sich die rechte Seite wie folgt definieren:

$$A_{ext}(w, z, \tau) = A_{mech} \sin(f\tau) + A_{Feld}(z + w) \quad (2.1.10)$$

Dabei bedeuten f eine normierte Frequenz, die sich als Verhältnis von Anregungsfrequenz und Eigenfrequenz darstellt und A_{Feld} die Wirkung des Kraftfelds. Für kleine Auslenkungen w - eine Situation, die quasi immer in AFM-Systemen vorliegt - kann durch Taylorentwicklung folgende Form erhalten werden:

$$A_{ext}(w, z, \tau) = A_{mech} \sin(f\tau) + A_{Feld}(z) + w \frac{\partial A_{Feld}(z)}{\partial z} \quad (2.1.11)$$

Einsetzen in (2.1.9) und Unterdrücken des statischen Terms liefert schließlich:

$$w'' + \frac{1}{Q} w' + \left(1 - \frac{\partial A_{Feld}(z)}{\partial z}\right) w = w'' + \frac{w'}{Q} + \alpha w = A_{mech} \sin(f\tau) \quad (2.1.12)$$

Die Lösung dieser typischen Gleichung ist allgemein bekannt, die wichtigen Parameter Amplitudengang und Phasengang können angegeben werden (siehe auch Abbildung 2.5):

$$A = \frac{1}{\sqrt{(\alpha - f^2)^2 + \frac{f^2}{Q^2}}}; \varphi = \arctan\left(\frac{f}{Q(\alpha - f^2)}\right); w(\tau) = A \sin(f\tau + \varphi) \quad (2.1.13)$$

Dabei wurden die Terme vor der Auslenkung w der Einfachheit halber zu einer neuen Konstante α zusammengefasst. Zur Messung in einem tatsächlichen Experiment bieten sich sowohl die Amplituden- als auch Phasenverschiebung relativ zu den Werten für $\alpha=1$ an:

$$\begin{aligned} \Delta A &= A(\alpha \neq 1) - A(\alpha = 1) \approx \frac{1 - f^2}{A(\alpha = 1)^3} \Delta \alpha \\ \Delta \varphi &= \varphi(\alpha \neq 1) - \varphi(\alpha = 1) \approx Q \Delta \alpha \end{aligned} \quad (2.1.14)$$

Obwohl beide Darstellungen in (2.1.14) Linearität in $\Delta \alpha$ suggerieren, gilt diese Annahme im Fall der Amplitude tatsächlich nur für sehr kleine Abweichungen (Abbildung 2.6).

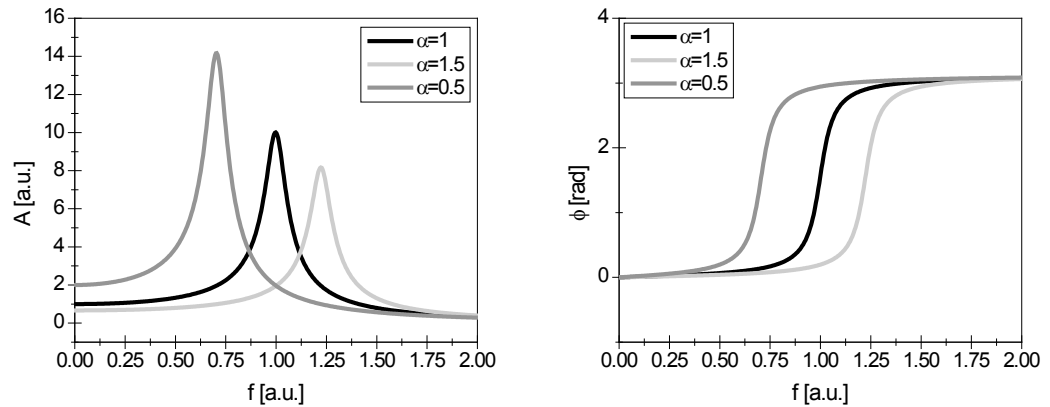


Abbildung 2.5: Amplituden- und Phasengang für Cantileverschwingungen in einem statischen Kraftfeld. Die Graphen wurden für einen festen Qualitätsfaktor $Q=10$ und für positive ($\alpha < 1$) und negative ($\alpha > 1$) Kraftgradienten generiert.

Allerdings können und werden beide Signale in Experimenten genutzt: Über die Amplitude wird der Abstand zwischen Spitze und Probe im Non-Contact-Modus geregelt, da diese stark von der auftretenden Kraft abhängt (Anhang B,[8]). Das Phasensignal kann entweder im selben Modus parallel zur Erlangung weiterer Informationen genutzt werden oder aber zur Abbildung der Gradienten kleiner zusätzlicher Kräfte, wie zum Beispiel der Coulombkraft oder kapazitiven Wechselwirkungen [9].

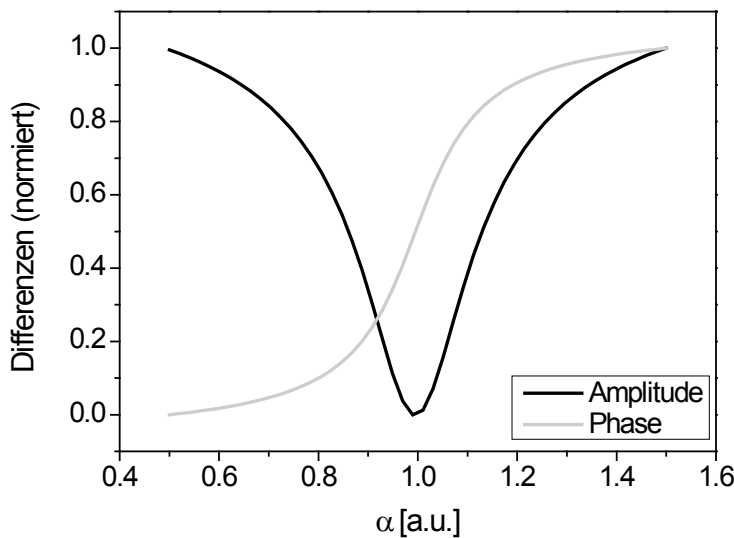


Abbildung 2.6: Änderung von Amplitude und Phase bei Variation des Parameters α . Offenbar reagiert das Phasensignal deutlich linearer auf diese Änderung.

2.) Mechanische Anregung in stufenförmigen Kraftfeld

Der Fall einer angeregten Schwingung des Cantilevers in einem stufenförmigen Potential, was eine zwar grobe aber nützliche Näherung des repulsiven Anteils der Spitze-Probe-Wechselwirkung darstellt, wird im sogenannten Semi-Contact-Modus (auch Tapping-Mode) genutzt. Dabei kann man annehmen, dass die Spitze nicht in die

Probe eindringen kann, die Schwingungsamplitude somit durch den Spitze-Probe-Abstand limitiert wird. Die zugehörige rechte Seite lautet (siehe Gleichung (2.1.10)):

$$A_{ext}(w, z, \tau) = A_{mech} \sin(f\tau) + \Theta(z + w) \quad (2.1.15)$$

In (2.1.15) steht Θ für eine Sprungfunktion. Analytische Lösungen sind nicht zugänglich, deshalb sollen einige wichtige numerische Lösungen für das Problem (2.1.9) & (2.1.15) angegeben werden. Sowohl der zeitliche Verlauf (Abbildung 2.7a) als auch der Amplitudengang sind maßgeblich durch den Spitze-Probe-Abstand z definiert. Genauer gesagt existiert unterhalb eines gewissen Schwellabstands ein linearer Zusammenhang (Abbildung 2.7b). Die in den Grafiken für verschiedenen Gütefaktoren Q erkennbaren Überschwinger stellen dabei die Auswirkung der Verschiebung der Resonanzfrequenz in der Nähe der Barriere, herleitbar ausgehend von (2.1.13), dar und sollen hier nicht näher erläutert werden. Wichtig ist das Vorhandensein einer linearen Regelgröße, in diesem Fall der Amplitude. Es bleibt zu erwähnen, dass Setups basierend auf diesem Prinzip die Vorteile von Contact- und Non-Contact-Modi vereinen, weshalb sie sehr häufig zur Topografieabbildung genutzt werden [10].

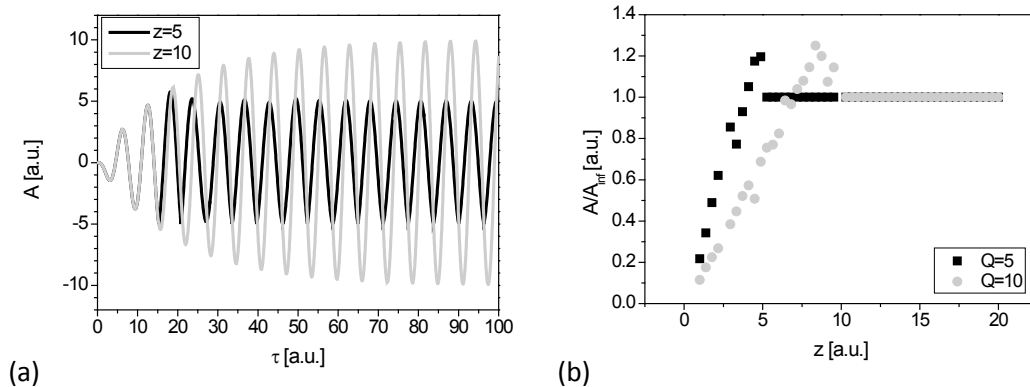


Abbildung 2.7: Zeitliches Verhalten der Cantileverschwingung in stufenförmigen Kraftfeld für verschiedene Höhen (a) und Amplitudengang in Abhängigkeit von dieser Höhe (b, Amplituden auf „freie“ Amplitude normiert). Offenbar existiert unterhalb eines minimalen Abstandes ein linearer Zusammenhang zwischen Amplitude und Abstand, wobei der Anstieg vom Gütefaktor des Systems abhängt.

3.) Elektrostatische Anregung bei kapazitiver Wechselwirkung

In diesem, für diese Arbeit ungleich wichtigsten Fall, ergibt sich folgende externe Anregungsamplitude:

$$A_{ext}(w, z, \tau) = C_{es}(z + w)(U - U_0) \sin(f\tau) \quad (2.1.16)$$

Die genaue Form der Abhängigkeit $C_{es}(z)$ soll später besprochen werden, für den Moment genügt die Annahme einer glatten Funktion, deren Variation bei Änderung der Koordinate z gering ist. Die grundlegendste Abhängigkeit hierbei wird durch den Term vermittelt, der die Spannung U und die, willkürliche, Referenzspannung U_0 enthält.

Augenscheinlich können die bereits ermittelten Formulierungen für den Fall übernommen werden, sodass die gemessene Amplitude proportional zur Differenz zwischen U und U_0 wird.

2.2 Das experimentelle Setup

Zusätzlich zu den grundlegenden Darstellungen bezüglich üblicher Rasterkraftmikroskope soll in diesem Abschnitt das tatsächlich genutzte System und die angewendeten Modi erläutert werden. Das genutzte Gerät trägt den Namen „Smena“ und wird von „NTMDT“ hergestellt sowie vertrieben. In der vorliegenden Konfiguration erlaubt es neben den Standard-Modi zur Abbildung der Oberfläche einer Probe auch die Messung elektrischer Eigenschaften. Diese beinhalten die mittlerweile weitverbreitete Anwendung „Conductive AFM“ (C-AFM), welche punktuelle Strommessungen ausnutzt, und die äußerst interessant erscheinende Möglichkeit, elektrostatische Wechselwirkungen abzubilden. Der verbaute Scanner ermöglicht Abtastbereiche von 60 μm lateral und 2 μm vertikal mit Auflösungen im Bereich von 10 nm und 0.1 nm. Zur Detektion wird ein streifenförmig fokussierter Diodenlaser (650 nm, etwa 1 mW) genutzt, wobei die beleuchtete Fläche (Größe des Laserstrahls) auf der Probe etwa 50 μm x 5 μm groß ist.

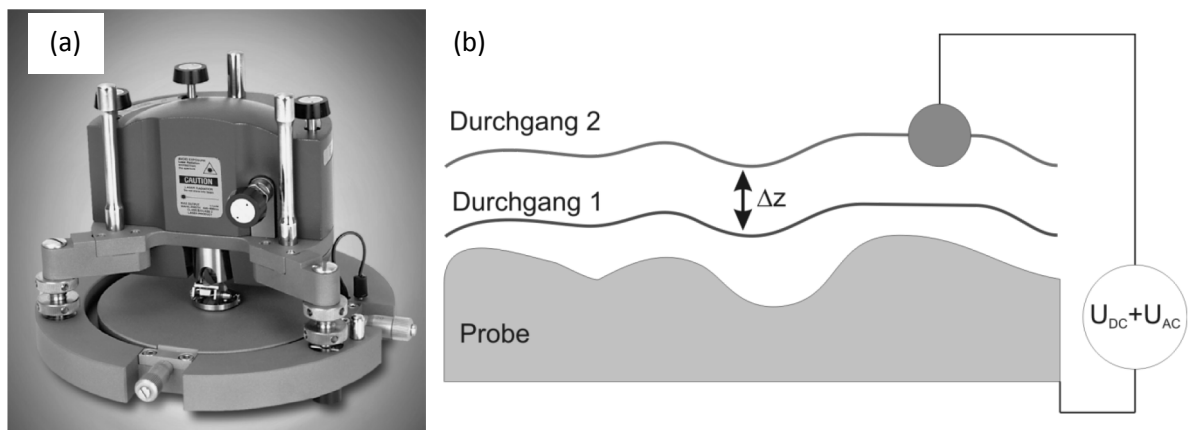


Abbildung 2.8: (a) Messkopf des verwendeten Smena-Systems. (b) Schematische Darstellung der zweiphasigen Messungen: Die als Kreis stilisierte Spitze des System wird auf einem Pfad mit konstantem Abstand Δz zur Oberfläche geführt und kann aufgrund angelegter Spannungen zur Messung elektrostatischer Wechselwirkungen genutzt werden.

Die der Arbeit zugrunde liegenden Messungen wurden zweiphasig durchgeführt (vergleiche Abbildung 2.8b), das heißt:

1.) Messung eines Höhenprofils

Dazu wurde eine Linie der Oberfläche der Probe im Tapping-Mode gescannt und das gemessene Profil zwischengespeichert.

2.) Messung elektrostatischer Wechselwirkung

Dieselbe Linie wurde vom Mikroskop erneut gescannt, allerdings ohne Abstandsregelung mittels mechanischer Wechselwirkung.

Vielmehr wurde das zuvor gespeicherte Profil genutzt, um einen definierten Abstand Δz zur Oberfläche einzuhalten⁴. Der Cantilever wird in diesem Durchgang verschiedenartig angeregt, was folgende prinzipielle Möglichkeiten erlaubt:

a) Mechanische Anregung und gleichzeitiges Anlegen einer Gleichspannung

In diesem Modus wird im zweiten Durchgang die - spannungsabhängige - Phasenverschiebung zwischen Erregerschwingung und Cantileverschwingung gemessen. Gemäß der Beschreibung im Abschnitt „Die Dynamik des Spitze-Probe-Systems“ (Seite 14) ergibt sich:

$$\Delta\varphi \approx \frac{Q}{k} \frac{\partial F_{ES}(U, z)}{\partial z} \quad (2.2.1)$$

Die detektierte Phasendifferenz hängt also neben den Kenngrößen des Systems, Gütefaktor Q und Federkonstante k , lediglich vom vertikalen Gradienten der elektrostatischen Kraftwirkung F_{ES} auf die Spitze ab.

b) Elektrische Anregung ohne Regelung

Die Oszillation des Cantilevers wird über eine angelegte Wechselspannung realisiert, die allgemein akzeptierte Formulierung für die wirkende Kraft ist [11]:

$$F(U_{AC}, U_{DC}, z) = K(z) (U_{DC} - U_0 + U_{AC} \sin(\omega t))^2 \quad (2.2.2)$$

Die die Geometrie vermittelnde Konstante K soll im Moment nicht näher erläutert werden. Offenbar lässt sich der Anteil der Kraft, der von der Gleichspannung U_{DC} , der anregenden Wechselspannung U_{AC} und einer Referenzspannung U_0 abhängt, umschreiben zu (mit der Vereinfachung $\Delta U = U_{DC} - U_0$):

$$\begin{aligned} \frac{F}{K} &= \Delta U^2 + 2\Delta U \cdot U_{AC} \sin(\omega t) + U_{AC}^2 \sin^2(\omega t) \\ \frac{F}{K} &= \Delta U^2 + \frac{1}{2} U_{AC}^2 + 2\Delta U \cdot U_{AC} \sin(\omega t) - \frac{1}{2} U_{AC}^2 \cos(2\omega t) \\ \frac{F}{K} &= A_0 + A_\omega + A_{2\omega} \end{aligned} \quad (2.2.3)$$

Es sind somit Komponenten enthalten, welche zeitlich konstant sind (A_0) oder mit der Anregungsfrequenz oder deren Doppelten schwingen (A_ω und $A_{2\omega}$). Nutzt man nun einen, in jedem AFM vorhandenen, phasenempfindlichen Detektor, können die entsprechenden Kraftanteile abgebildet werden. Man erhält Auskunft entweder über die kapazitiven Eigenschaften (doppelte Frequenz) oder, bei lateral nahezu konstanter Kapazität, über Spannungsverteilungen.

c) Elektrische Anregung mit Regelung

⁴ Der einstellbare Parameter ist ein vertikaler Abstand zur mittleren Position der Spitze im ersten Durchgang, die tatsächliche Höhe ergibt sich also als Summe von diesem Wert und der Schwingungsamplitude im ersten Durchgang.

Nutzt man den Anteil der Kraft, welcher mit der Anregungsfrequenz schwingt, um einen Regelkreis zu betreiben, ergibt sich Folgendes: Da augenscheinlich die Kraftkomponente und damit die gemessene Amplitude für $\Delta U=0$ verschwindet, kann die punktweise Abbildung der dazu nötigen Spannung U_{DC} als Karte der Spannungsverteilung auf der Oberfläche verstanden werden. Diese Variante wird allgemein als Scanning Kelvin Probe Microscopy (SKM) bezeichnet und hat für die Beurteilung von Oberflächen auf einer nm-Skala einen sehr hohen Stellenwert.

Es existieren in der Literatur keine üblichen Bezeichnungen für die in den Punkten a und b beschriebenen Methoden, im Rahmen dieser Arbeit soll folgende Nomenklatur genutzt werden: Da Methode b hauptsächlich kapazitive Eigenschaften abbildet, wird hierfür der Begriff nc-SCM (noncontact Scanning Capacitance Microscopy) benutzt, wobei das Präfix meist durch die Angabe der Frequenz ersetzt wird (zum Beispiel 2ω -SCM).

Zur Illustration der Ergebnisse der verschiedenen Methoden soll folgendes, stark vereinfachtes, Gedankenexperiment genutzt werden: Die Spitze des Systems soll für den Moment als Punktladung q modelliert werden, damit gilt:

$$\begin{aligned} F(x, y, z) &= q(U_{DC})E(z) = -q(U_{DC})\frac{\partial V(x, y, z)}{\partial z} \\ \frac{\partial F(z)}{\partial z} &= -q(U_{DC})\frac{\partial^2 V(x, y, z)}{\partial z^2} \end{aligned} \quad (2.2.4)$$

Dabei beschreibt $V(x, y, z)$ die Potentialverteilung, die sich im oberen Halbraum $z>0$ bei bekannter Verteilung des Potentials auf der Ebenen $z=0$ einstellt. Für diesen Fall liefert die SKM-Methode offenbar den Wert des Potentials V am Ort (x, y, z) , die ω -SCM-Methode einen Wert, der proportional zum elektrischen Feld ist, und die in (a) beschriebene Methode ein Signal proportional zur zweiten Ableitung des Potentials V . Es genügt also zur Visualisierung Spuren von V , dV/dz und d^2V/dz^2 zu generieren. Dazu wird eine sehr einfache Geometrie genutzt, eine streifenförmige Potentialverteilung der Breite $2w$ in x-Richtung (Abbildung 2.9).

Mathematisch wird das Problem durch die Anwendung der Poissonschen Integralformel lösbar ([12], Kapitel 4.2.2), es gilt:

$$V(x, y, z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} \frac{z}{\left(z^2 + (x-a)^2 + (y-b)^2\right)^{\frac{3}{2}}} V_0(a, b) da db \quad (2.2.5)$$

In (2.2.5) beschreibt V_0 die Potentialverteilung in der Ebene.

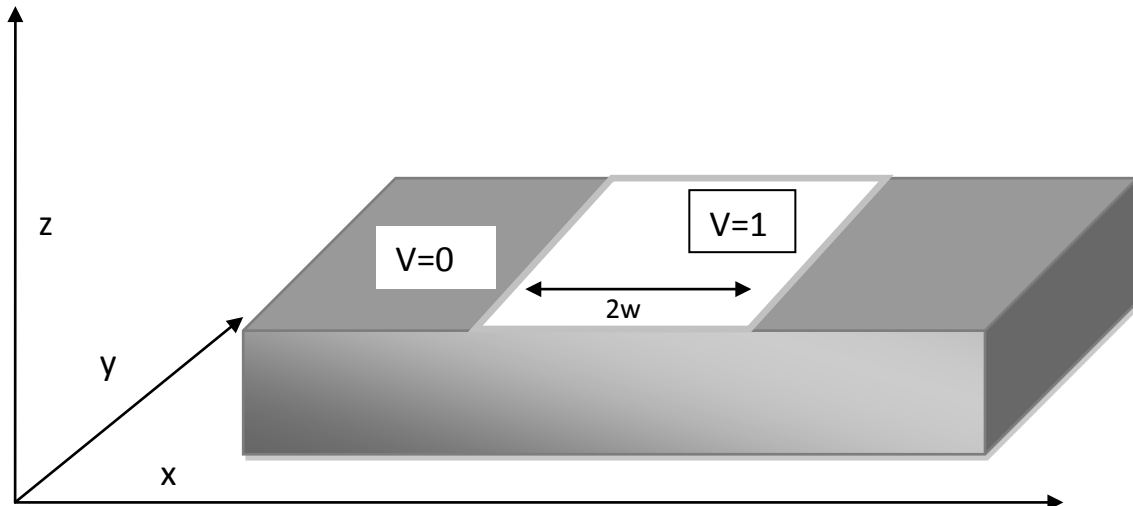


Abbildung 2.9: Geometrie zur Veranschaulichung des Informationsgehaltes der elektrostatischen Messungen

Für die hier unendlich angenommen Ausdehnung des Streifens in y-Richtung ergibt sich weiter:

$$V(x, z) = \int_{-\infty}^{\infty} \frac{1}{\pi} \frac{z}{z^2 + (x-a)^2} V_0(a) da = \frac{1}{\pi} \left(\arctan\left(\frac{x+w}{z}\right) - \arctan\left(\frac{x-w}{z}\right) \right)$$

$$\frac{\partial V(x, z)}{\partial z} = \frac{2w}{\pi} \frac{z^2 - x^2 + w^2}{[z^2 + (x-w)^2][z^2 + (x+w)^2]} \quad (2.2.6)$$

$$\frac{\partial^2 V(x, z)}{\partial z^2} = \frac{4wz(z^4 - 2x^2z^2 + 2w^2z^2 - 3x^4 + 2w^2x^2 + w^4)}{\pi[z^2 + (x-w)^2]^2[z^2 + (x+w)^2]^2}$$

Die so gewonnenen analytischen Ausdrücke sind in Abbildung 2.10 für eine Höhe von $z=1$ und eine Weite $w=1$ dargestellt.

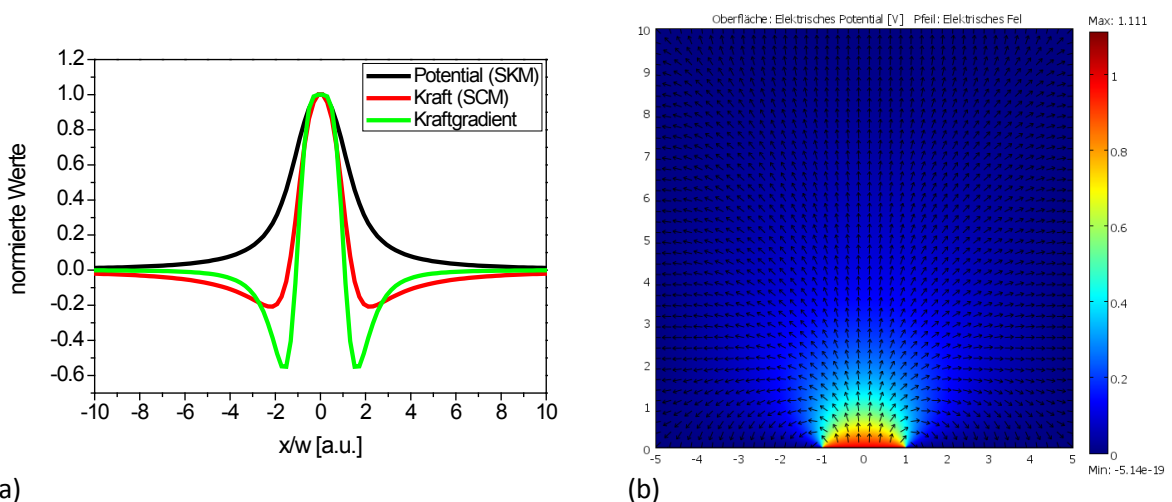


Abbildung 2.10: (a) Auswertung der Ergebnisse von (2.2.6) für $z=w=1$. Offenbar sind alle besprochenen Methoden zur Abbildung von Potentialverteilungen geeignet, allerdings mit Unterschieden. Während SKM tatsächlich ein Abbild des Potentials liefert, zeigen die anderen Ansätze zusätzliche laterale Einflüsse, was eine Quantifizierung erschwert, wenn nicht sogar unmöglich macht. (b) Numerische Lösung des Problems, welche die Krümmung der Feldlinien veranschaulicht. Diese sind für die Vorzeichenwechsel in Abbildung (a) verantwortlich.

Alle benutzten Methoden erlauben eine zufriedenstellende Abbildung des Potentialverlaufs, im Falle der Abbildung der Gradienten ergeben sich allerdings zusätzliche laterale Effekte, welche sich auf die Krümmung der Feldlinien zurückführen lässt. Diese erzeugt sowohl einen Vorzeichenwechsel in der Kraft als auch im Kraftgradienten.

Obwohl, im Moment lediglich gestützt auf die Ergebnisse des doch sehr einfach gewählten Modells, die Quantifizierung der unregelmäßigen kraftabbildenden Methoden nicht möglich ist, scheinen sie allerdings bestens geeignet, um Kontraste zu erzeugen. Falls also zum Beispiel schnell die Position einer Störung bestimmt werden soll, können diese Aufbauten zum Einsatz kommen.

Abschließend soll an dieser Stelle noch geklärt werden, warum der zweiphasige Messaufbau genutzt wurde, obwohl andere, simultane, Ansätze verfügbar sind. Grundlage dieser alternativen Methoden ist immer eine gleichzeitige elektrische und mechanische Anregung des Cantilevers mit unterschiedlichen Frequenzen, wobei dann durch Nutzung frequenzselektiver Messtechnik die Signale zur Topografiebestimmung und Messung elektrostatischer Phänomene getrennt bearbeitet werden. Die tatsächliche Bestimmung zum Beispiel des Oberflächenpotentials geschieht dabei entweder durch Auswertung von Amplituden- oder Phasensignalen [11], [13], welche sich etwas in ihrem lateralen Auflösungsvermögen unterscheiden. Die Befürworter der Einphasentechniken nennen zwei Argumente:

1.) Geringerer Zeitaufwand

Dadurch, dass Topografie und elektrisches Signal in einem Durchgang gemessen werden kann, halbiert sich offenbar der Zeitaufwand. Gegen dieses Argument ist nichts einzuwenden, allerdings ist der erhöhte instrumentelle Aufwand zu beachten, welcher im Endeffekt die Vorteile möglicherweise relativiert.

2.) Höhere laterale Auflösung gegenüber zweiphasigen Messungen

Weithin verbreitet ist der Glaube, dass das üblicherweise genutzte vertikale Offset zwischen den zwei Durchgängen eine Verschlechterung der Auflösung mit sich bringt. Diese Auffassung ist nur bedingt richtig: Geht man zum Beispiel von der in (2.2.6) genutzten Kernelfunktion für den Fall quasi zweidimensionaler Verteilung aus, was anschaulich der Messung mit einer punktförmigen Sonde entspricht, erkennt man, dass das Messsignal eine Faltung eben jener Verteilung mit einer Lorentzfunktion ist. Das resultiert in einem Aufweichen von Kanten, speziell werden Stufen in Arcustangensfunktionen gewandelt. Die Breite der Lorentzfunktionen, und damit auch die Weite der Aufweichung, entspricht der Höhe des Messpunktes über der Ebene, was das Argument der schlechteren Auflösung motiviert. Tatsächlich ist es aber so, dass durchaus auch negative Offsets benutzt werden können, was eine Annäherung an die Oberfläche und somit verbesserte Auflösung bedeutet.

Nach der Entkräftung der Nachteile des genutzten Aufbaus kann auch ein deutlicher Vorteil genannt werden: Es gibt keine Kopplung zwischen Topografiemessung und Potentialbestimmung. Die simultane Messung kann wie folgt beschrieben werden (vgl. Kapitel Rastersondenmikroskopie):

Konzeptionelle Grundlagen

$$w'' + \frac{1}{Q}w' + w = A_{mech} \sin(f\tau) + A_{ES,F} \sin(F\tau) + A_{LJ}(z+w) + A_{ES,0}(z+w) \quad (2.2.7)$$

Es existiert somit neben dem statischen Kraftfeld, welches durch atomare Kräfte zwischen Spitze und Probe verursacht wird (A_{LJ}), ein zusätzliches Feld aufgrund elektrostatischer Wechselwirkung ($A_{ES,0}$). Für kleine Schwingungsamplituden kann wieder um z entwickelt werden, was bei Unterdrückung der statischen Anteile zu folgendem Ausdruck führt:

$$w'' + \frac{1}{Q}w' + \left(1 - \frac{\partial A_{LJ}}{\partial z} - \frac{\partial A_{ES,0}}{\partial z}\right)w = A_{mech} \sin(f\tau) + A_{ES,F} \sin(F\tau) \quad (2.2.8)$$

Die Lösung ergibt sich wie erwartet als Überlagerung zweier Anteile unterschiedlicher Frequenz, allerdings werden die Amplituden beider Schwingungen durch beide Kraftfelder beeinflusst:

$$A_f = \frac{A_{mech}}{\sqrt{\left(1 - \frac{\partial A_{LJ}}{\partial z} - \frac{\partial A_{ES,0}}{\partial z} - f^2\right)^2 + \frac{f^2}{Q^2}}} \quad (2.2.9)$$

Da das Regelsystem nun versucht, die Amplitude A_f der Schwingung mit der Frequenz der mechanischen Anregung konstant zu halten, indem der Spitze-Probe-Abstand variiert wird, kann es zu schlecht kompensierbaren Beeinflussungen kommen. Wird, wie üblich, zur Anregung die Resonanzfrequenz des Cantilevers für große Spitze-Probe-Abstände verwendet, ergibt sich für den Amplitudengang bei Zusammenfassung der zwei auftretenden Gradienten zu einer Funktion G :

$$\frac{A_f}{A_{mech}} = \frac{1}{\sqrt{G(z)^2 + \frac{1-G(z)}{Q^2} - \frac{1}{4Q^4}}} \quad (2.2.10)$$

Ausgehend von Abschätzungen für den Verlauf der Kraft-Abstands-Funktionen (Anhang B, Modell homogen geladene Fläche-Punktladung) ergibt sich für $G(z)$ zum Beispiel:

$$G(z) = -\frac{C_1}{z^3} + \frac{C_2}{z^2} \quad (2.2.11)$$

Einsetzen dieser Funktion mit entsprechenden Parametern C_1 und C_2 generiert Amplitudengänge wie in Abbildung 2.11 dargestellt. Im Normalfall, also ohne zusätzliche elektrostatische Kräfte, bewegt sich das System auf der schwarzen Spur, eine Abstandsverringerung führt zu einer Amplitudenabnahme, was das System wiederum durch Abstandsvergrößerung kompensiert. Falls nun aber aufgrund von Potentialdifferenzen zwischen Spitze und Probe eine Amplitudenabnahme verursacht wird (Übergang zu einer der anderen Spuren), ist der anschließende Regelvorgang offenbar falsch, da keine Höhenvariation vorlag. Somit kann es im ungünstigsten Fall vorkommen, dass glatte Oberflächen als strukturiert abgebildet werden und, durch die simultane Messung, der Verlauf des Oberflächenpotentials obstruiert wird.

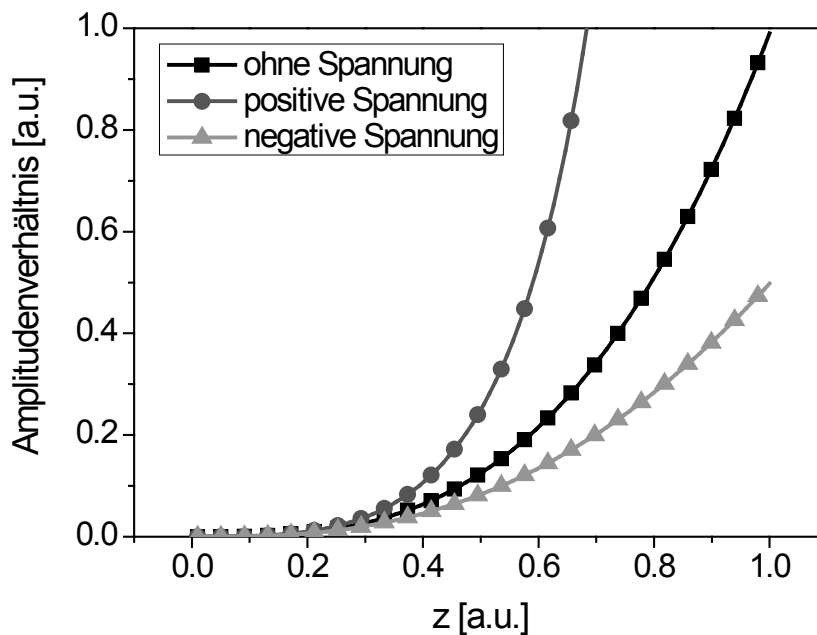


Abbildung 2.11: Abgeschätzte Amplitudengänge für simultane Messung von Topographie und elektrischem Potential. Die elektrostatische Wechselwirkung bewirkt eine zusätzlich Verschiebung der Resonanzfrequenz, was unterschiedliche Amplituden verursacht.

Diese Probleme wurden durch Nutzung der zweiphasigen Methode fast ausgeschlossen, da der genutzte Tapping-Mode weniger anfällig für externe Kräfte ist.

2.3 Numerische Behandlung von Randwertproblemen

Physikalische Problemstellungen, beziehungsweise deren mathematische Formulierung, können üblicherweise in zwei Klassen eingeteilt werden: Anfangswertprobleme (AWPs beziehungsweise IVPs) und Randwertprobleme (RWPs beziehungsweise BVPs).

2.3.1 Anfangswertprobleme

Eine typische Darstellung eines AWP ist:

$$\begin{aligned} \frac{\partial}{\partial t} y(t) &= f(y(t), t) \\ y(t_0) &= y_0 \end{aligned} \tag{2.3.1}$$

Hierbei sind y die gesuchte Lösung, f deren Ableitung, t die Zeit (oder ein äquivalenter Parameter) und y_0 der Wert der gesuchten Lösung am Anfangszeitpunkt t_0 .

Ein beliebtes, und auch einfaches, Beispiel für ein Anfangswertproblem ist der gedämpfte harmonische Oszillator. Nach Umformulierung der Differentialgleichung zweiter Ordnung (die Gleichung verknüpft die erste und die zweite Ableitung der Funktion mit dieser) in ein System erster Ordnung kann das Problem numerisch behandelt werden. Die Lösung ergibt sich durch Integration der Bewegungsgleichungen (γ ist der Dämpfungskoeffizient, ω_0 die Eigenfrequenz des Schwingers und $v(t)$ die Geschwindigkeit):

$$\begin{pmatrix} y(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} y_0 \\ v_0 \end{pmatrix} + \int_{t_0}^t \begin{pmatrix} v(x) \\ -2\gamma v(t) - \omega_0^2 y(x) \end{pmatrix} dx \quad (2.3.2)$$

Die numerische Behandlung reduziert sich somit auf eine Integration, wofür verschiedene Methoden zur Verfügung stehen. Meist werden Einschritt-Verfahren benutzt, häufig einfache Euler- oder „Runge-Kutta“-Löser, auf deren genauere Beschreibung hier verzichtet werden soll (nähere Erläuterungen zu den Verfahren finden sich in [14] und [15]). Diese Algorithmen erzeugen brauchbare Lösungen nur, wenn die gewählten Schrittweiten im Bereich einiger Hundertstel der Zeitkonstante des Systems liegen. Eine oft genutzte Möglichkeit, um das Problem der auftretenden numerischen Dämpfung zu umgehen, ist die Anwendung adaptiver Löser, die ihre Schrittweite dem jeweiligen lokalen Fehler anpassen.

Abbildung 2.12 zeigt einen Vergleich der numerischen Fehler, hier definiert als Varianz der relativen Fehler der Datenpunkte, für verschiedene Abtastraten und Algorithmen: Während der Fehler für den adaptiven „RK45“-Löser fast konstant ist, ist eine starke Abhängigkeit von der Abtastrate für den Euler-Löser zu erkennen.

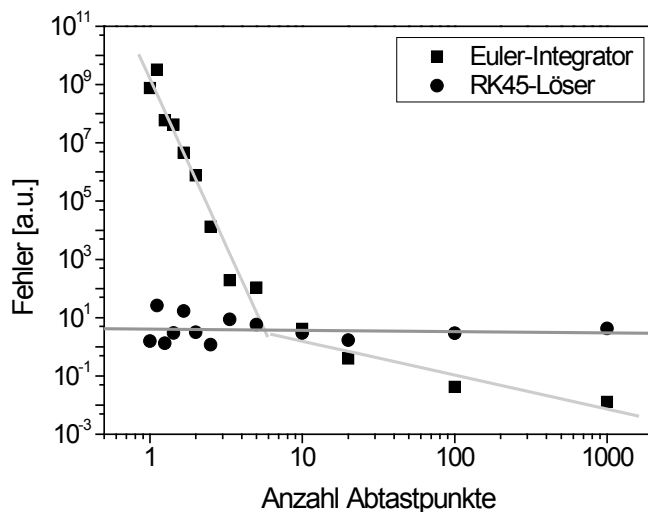


Abbildung 2.12: Vergleich der Fehler für Euler- und RK45-Löser. Die in dieser Darstellung linearen Anpassfunktionen verdeutlichen die exponentielle Abhängigkeit des Fehlers bei Verwendung konstanter Schrittweite (Quadrate), während der Fehler bei Nutzung adaptiver Methoden (Kreise) fast konstant bleibt.

2.3.2 Randwertprobleme

Ein Anfangswertproblem kann zu einem Randwertproblem werden, wenn nicht für alle Freiheitsgrade Anfangswerte gegeben sind, allerdings Endwerte gefordert werden. So zum Beispiel könnte man bei der besprochenen Schwingungsgleichung die anfängliche Auslenkung und die Auslenkung nach einer gewissen Zeit vorschreiben. Da für eine direkte Integration allerdings auch die anfängliche Geschwindigkeit notwendig wäre, muss eine Lösung anderweitig konstruiert werden.

Das abgewandelte Problem lautet also nunmehr:

$$\begin{aligned}
\frac{\partial}{\partial t} \begin{pmatrix} y(t) \\ v(t) \end{pmatrix} &= \begin{pmatrix} v(t) \\ -2\gamma v(t) - \omega_0^2 y(t) \end{pmatrix} \\
y(0) &= y_0 \\
y(t_E) &= y_E
\end{aligned} \tag{2.3.3}$$

Die „Shooting“-Methode

Einen intuitiven Algorithmus nutzt die „Shooting“-Methode: Die Endwerte des zu lösenden System werden als Funktion der freien Anfangsparameter betrachtet und es wird der Satz dieser Anfangsparameter gesucht, der die geforderten Endwerte erzeugt. Die gesuchte Lösung wird demnach durch mehrfache Integration und Nullstellensuche bestimmt. Für das benutzte RWP der Schwingungsgleichung ergibt sich folgende Vorschrift:

1. Löse das AWP mit den Anfangswerten y_0 und v
2. Werte die Lösung an der rechten Intervallgrenze aus
3. Benutze die Differenz des errechneten Werts und des geforderten Werts als rechte Seite einer Nullstellensuche
4. Wiederhole den Ablauf mit dem generierten, neuen, v

$$\begin{aligned}
RWP: \frac{\partial^2}{\partial t^2} y(t) &= f\left(\frac{\partial}{\partial t} y(t), y(t), t\right), y(t_0) = y_0, y(t_E) = y_E \\
AWP: \frac{\partial^2}{\partial t^2} y(t) &= f\left(\frac{\partial}{\partial t} y(t), y(t), t\right), y(t_0) = y_0, \frac{\partial}{\partial t} y(t_0) = v \\
F(v) &= y(t_E, v) - y_E \stackrel{!}{=} 0
\end{aligned} \tag{2.3.4}$$

Wie der Abbildung 2.13 zu entnehmen ist, sinkt der Fehler im rechten Randwert exponentiell mit der Iteration, was demnach brauchbare Ergebnisse in annehmbarer Zeit bedeutet.

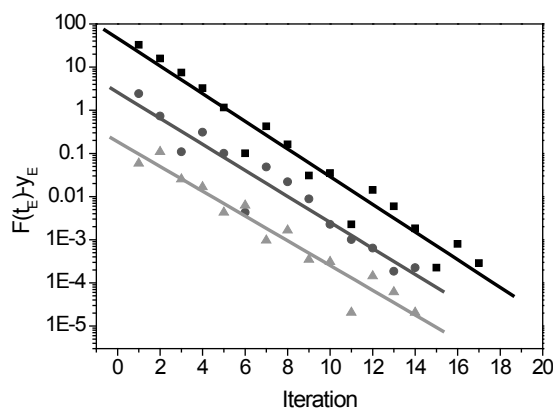


Abbildung 2.13: Entwicklung des Fehlers für das im Text besprochene Problem mit verschiedenen Startparametern, hier der Ableitung am linken Rand. Offenbar sinkt der Fehler jeweils exponentiell mit der Iterationszahl.

Unglücklicherweise stellt sich heraus, dass die Nutzung der Integration der Anfangsbedingungen bei bestimmten Problemstellungen Instabilitäten erzeugt. Als Beispiel soll hier die Poissongleichung in einem Halbleiter genutzt werden, welche nach Reskalierung (siehe Anhang A) folgende Gestalt annimmt:

$$\frac{\partial^2 u}{\partial x^2} = \sinh(u - u_F) + \sinh(u_F) \quad (2.3.5)$$

Darin sind u das dimensionslose Potential (relativ zur thermischen Spannung kT/e , etwa 25 mV bei Raumtemperatur), x der dimensionslose Ort (relativ zur intrinsischen Debyelänge in Silizium, etwa 22 μm) und u_F ein dimensionsloses Referenzpotential, das von der Dotierung abhängt. Ein typisches Randwertproblem wäre:

$$u(0) = 1, u(\infty) = 0 \quad (2.3.6)$$

Die physikalische Interpretation von (2.3.6) ist die Berechnung der Ladungsträgerverteilung in der Nähe der Flachbandbedingung, also für verschwindendes Potential an der Oberfläche. Die Stabilität der numerischen Behandlung soll an einer etwas abgewandelten Version der korrekten Gleichung (2.3.6) illustriert werden: der rechte Rand wird auf eine Position verschoben, die etwa 10 % der intrinsischen Debyelänge entspricht⁵. Die Ortsabhängigkeit der durch Integration erhaltenen Lösungen für verschiedene Anfangswerte des dimensionslosen elektrischen Feldes zeigt Abbildung 2.14. Wie im linken Plot dargestellt, ergibt sich lediglich eine stabile Lösung für einen Startwert in der Nähe des tatsächlichen, analytisch zugänglichen, Wertes von etwa -90. Die anderen gezeigten Spuren verlaufen nicht bis zum Ende des Intervalls, da die Berechnung wegen des Auftretens undefinierter Werte („Inf“) abgebrochen wurde. Dieses Verhalten ist auch im rechten Plot zu erkennen: die Endpunkte (schwarz) und auch die Endwerte (grau) liegen nur für eine einzige Parameterwahl, die richtige, auf den geforderten Werten.

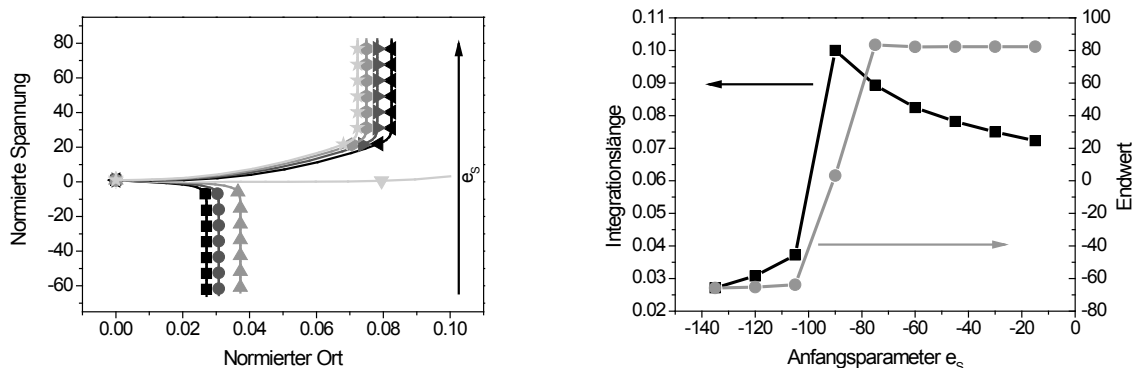


Abbildung 2.14: Abhängigkeit der numerischen Lösungen der nichtlinearen Poisson-Gleichung in Silizium für verschiedene Werte des elektrischen Feldes e_s an der Oberfläche. Nur in der Nähe des „richtigen“ Wertes (etwa -90) kann der gewählte Integrator die Berechnung komplett durchführen; für alle anderen Parameter ergeben sich nicht definierte Werte der Variablen.

Im Allgemeinen stellt das Auftauchen kleinerer numerischer Instabilitäten kein großes Problem dar, allerdings ist man in dem beschriebenen Fall, der Lösung des RWP mittels der „Shooting“-Methode, bei der Wahl des Algorithmus zur Nullstellensuche stark eingeschränkt: schnell konvergierende Methoden, wie das Newton-Verfahren, können nicht genutzt werden.

⁵ Eine Näherung für (2.3.5) ist $u'' = \cosh(u_F) \cdot u$ mit der Lösung: $u(x) = u(0) \cdot e^{-\frac{x}{\sqrt{\cosh(u_F)}}}$. Die Abklinglänge ist für typische Dotierungen, also übliche Werte von u_F , in der Größenordnung von 1% der intrinsischen Debyelänge, was die vorgenommene Anpassung motiviert.

Stattdessen muss ein langsames, stabileres Verfahren implementiert werden, zum Beispiel die Bisektionsmethode.

In Abbildung 2.15 sind die Ergebnisse des genannten Vorgehens für eine parametrische Studie gezeigt: das dimensionslose elektrische Feld (gleichzeitig der freie Parameter der Shooting-Methode) wurde in Abhängigkeit der dimensionslosen Spannung berechnet. Gleichzeitig wurde der Einfluss der Löserwahl und der Toleranzgrenze der Löser untersucht. Wie Abbildung 2.15a zu entnehmen ist, zeigen die gewählten Integratoren, natürlich abhängig von der gewählten Toleranz, zufriedenstellende Übereinstimmung mit der analytischen Lösung (siehe Anhang A). Allerdings sind sowohl Stabilität als auch Zeitaufwand sehr unterschiedlich, wie Abbildung 2.15b zeigt: Während der Fehler bei Wahl des „ode45“-Löser mit der Toleranz über mehrere Größenordnungen abnimmt, bleibt er beim „ode113“-Löser fast konstant. Des Weiteren ergibt sich, dass der erforderliche Zeitaufwand sich um den Faktor 10 unterscheidet, da „ode113“ schon bei geringen Toleranzen präzise arbeitet.

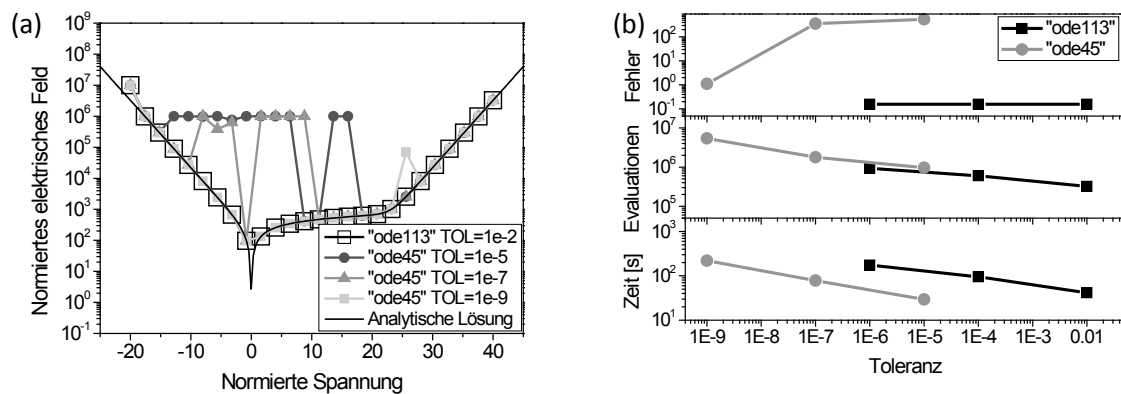


Abbildung 2.15: Mittels Bisektionsverfahren erhaltene Abhängigkeit des dimensionslosen elektrischen Feldes von der dimensionslosen Oberflächenspannung für verschiedene Integratoren und Fehlertoleranzen. Die Übereinstimmung der Ergebnisse mit der analytisch zugänglichen Lösung (schwarze Linie in a) ist, bei geeigneter Wahl der Parameter, gut. Es ergibt sich weiterhin, dass das Aufwand/Nutzen-Verhältnis stark variiert: der „ode113“-Integrator erweist sich bereits bei geringen Anforderungen als sehr genau (b).

Die Methode der finiten Differenzen

Eine weitere Klasse von Methoden zur näherungsweisen Lösung von Randwertproblemen wird als Methode der finiten Differenzen bezeichnet. Dabei werden die in den Gleichungen auftretenden Ableitungen durch diskrete Repräsentationen ersetzt. Dieses Vorgehen soll im Folgenden für die bereits diskutierte Poissongleichung besprochen werden.

Das zu untersuchende Ortsintervall L wird, wie in Abbildung 2.16 illustriert, in äquidistante Intervalle zerteilt:

$$[0, L] \rightarrow \{x_1 = 0, x_2 = x_1 + h, \dots, x_N = L\} \quad (2.3.7)$$

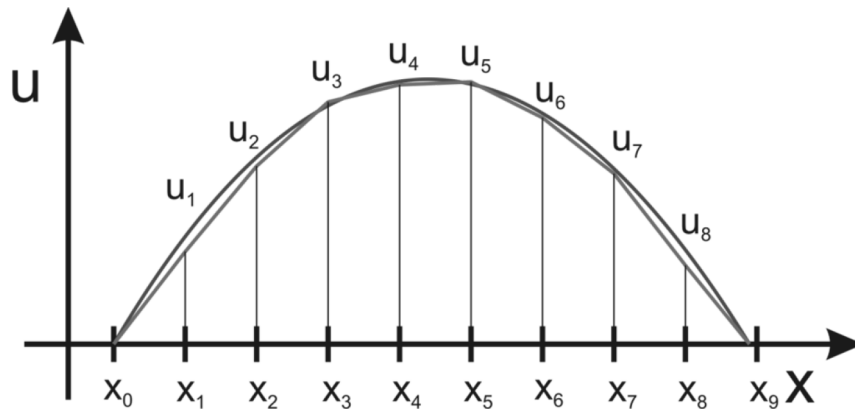


Abbildung 2.16: Beispiel einer näherungsweisen Lösung (grau) gewonnen durch die Methode der finiten Elemente. Die Werte x_i stellen dabei die Stützstellen dar und die Werte u_i die diskreten Funktionswerte an diesen Stützstellen.

Zu den Orten x_i gehören Werte u_i , was es erlaubt, die Differentialgleichung wie folgt zu formulieren:

$$\begin{aligned} \mathbf{L}(x_1, \dots, x_N, u_1, \dots, u_N) &= \mathbf{F}(u_1, \dots, u_N) \\ \forall i \in [2, N-1]: \frac{1}{h^2}(u_{i-1} - 2u_i + u_{i+1}) &= \sinh(u_i - u_F) - \sinh(u_F) \end{aligned} \quad (2.3.8)$$

Dabei stehen \mathbf{L} und \mathbf{F} für Operatoren, die auf den Vektoren der Orte und Lösungen definiert sind und den Laplace-Operator und die rechte Seite repräsentieren. Da die Schrittweite als konstant zu h angenommen wurde, kann die zweite Ableitung, wie angegeben, für alle Indizes i genähert werden⁶. Zusammengefasst ergibt sich somit ein Gleichungssystem der folgenden Struktur für die inneren Punkte:

$$\begin{pmatrix} -2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{N-2} \\ u_{N-1} \end{pmatrix} = h^2 \begin{pmatrix} \sinh(u_1 - u_F) + \sinh(u_F) \\ \sinh(u_2 - u_F) + \sinh(u_F) \\ \sinh(u_3 - u_F) + \sinh(u_F) \\ \vdots \\ \sinh(u_{N-2} - u_F) + \sinh(u_F) \\ \sinh(u_{N-1} - u_F) + \sinh(u_F) \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \quad (2.3.9)$$

Der zweite Vektor der rechten Seite repräsentiert die Randbedingungen, welche im ersten und letzten Wert stehen (hier also „1“ und „0“). In etwas angenehmerer Form lautet die zu lösende Aufgabe somit:

$$\mathbf{L}\vec{u} = h^2 \vec{f} - \vec{r} \quad (2.3.10)$$

Für den deutlich einfacheren Fall einer nicht vom Wert der Funktion abhängenden rechten Seite, ein Beispiel wäre die Modellierung einer geladenen Isolatorschicht zwischen zwei Metallplatten, ist die Lösung damit durch simple Multiplikation mit der Inversen der \mathbf{L} -Matrix zu finden:

⁶ Eine Taylorentwicklung der unbekannten Lösungsfunktion u um x_i liefert:

$$u(x_{i-1}) - 2u(x_i) + u(x_{i+1}) \approx u(x_i) - h \cdot u'(x_i) + \frac{h^2}{2} u''(x_i) - 2u(x_i) + u(x_i) + h \cdot u'(x_i) + \frac{h^2}{2} u''(x_i) = h^2 u''(x_i)$$

$$\vec{u} = \mathbf{L}^{-1} \left(-h^2 \frac{\rho}{\varepsilon} \vec{1} - \vec{r} \right) \quad (2.3.11)$$

In (2.3.11) ist dabei ρ die Ladungsdichte und $\vec{1}$ ein Vektor, der lediglich aus „1“ besteht. Zur Veranschaulichung soll hier $\rho/\varepsilon=1$ gesetzt werden und die Randwerte werden homogenisiert, das heißt an beiden Endpunkten des Intervalls sind die Randwerte „0“. Aufgrund der Linearität des Operators und der rechten Seite könnten beliebige Randwerte durch Addition einer linearen Funktion, welche die Randbedingungen erfüllt, erreicht werden.

Die Berechnung benötigt, bei Benutzung optimierter Algorithmen wie zum Beispiel „Umfpack“ ([16], enthalten in „Matlab“), auf aktuellen Systemen weniger als 10 μs pro Stützstelle, wie folgende Abbildung 2.17 illustriert.

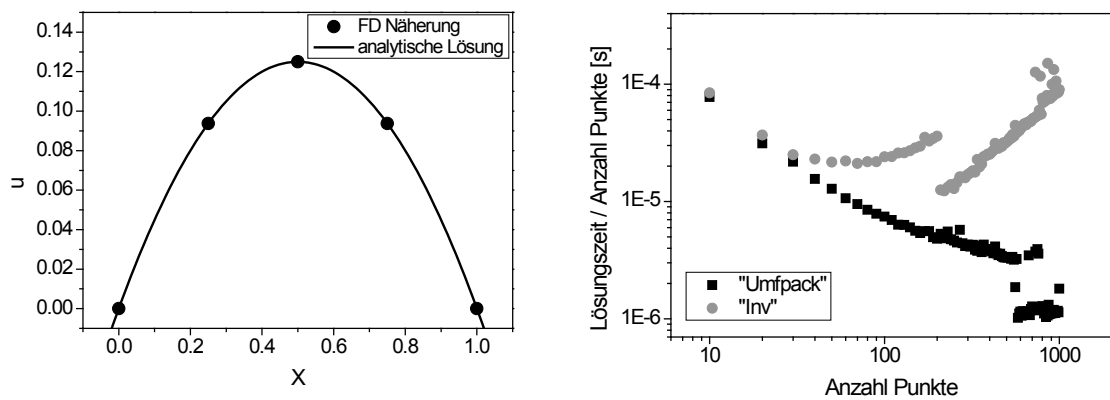


Abbildung 2.17: Lösung eines einfachen Poissonproblems mit Hilfe der Methode der finiten Differenzen. Die Lösung wird an diskreten Stützstellen berechnet (links). Der Zeitaufwand pro Stützstelle bleibt dabei etwa konstant in der Größenordnung von 10 μs (rechts). Es ist allerdings darauf hinzuweisen, dass bei großer Anzahl der Variablen die Effizienz, hier der zeitliche Aufwand, verschiedener Löser stark variiert.

Problematisch wird die Anwendung der finiten Differenzen bei Randbedingungen vom „Neumann“-Typ, also mit Vorgabe der Ableitung auf dem Rand. Während die Implementierung keine größere Mühe bereitet, kann das mathematische Gerüst zusammenbrechen, wie folgendes Beispiel verdeutlicht: Nimmt man die Ladungsdichte als stückweise konstant an und lässt das Vorzeichen wechseln, ergibt sich zum Beispiel folgende Problemstellung:

$$\begin{aligned} -\Delta u &= \begin{cases} 1; & x < 0 \\ -1; & x > 0 \end{cases} \\ \frac{du}{dx} \Big|_{x=-1} &= \frac{du}{dx} \Big|_{x=1} = 0 \end{aligned} \quad (2.3.12)$$

Gleichung (2.3.12) beschreibt zum Beispiel den Verlauf des Potentials an einem p-n-Übergang im Rahmen der Verarmungsnaheherung.

Die geforderten Randbedingungen können zum Beispiel durch Hinzunahme von „Geisterpunkten“ außerhalb des Intervalls und Auswertung des zentralen Differenzenquotienten geschehen:

$$\left. \frac{du}{dx} \right|_{x=0} \approx \frac{u_1 - u_{-1}}{2h} = 0 \Rightarrow u_{-1} - 2u_0 + u_1 = -2u_0 + 2u_1 \quad (2.3.13)$$

Die **L**-Matrix hat für 5 Punkte somit folgende, leicht modifizierte, Gestalt:

$$\mathbf{L} = \begin{pmatrix} -2 & 2 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 2 & -2 \end{pmatrix} \quad (2.3.14)$$

Unglücklicherweise ist diese Matrix singulär, das heißt, eine Lösung des Gleichungssystems ist nicht direkt möglich. Abhilfe schaffen hier sogenannte iterative Lösungsverfahren, zum Beispiel die Methode der konjugierten Gradienten [17].

Ziel der Methode ist wiederum die Lösung eines Gleichungssystems der Form $\mathbf{A}x = b$, wobei **A** eine symmetrische, positiv definite Matrix ist, x ein Lösungsvektor und b eine vorgegebene rechte Seite. Die entscheidende Idee ist, den Lösungsraum durch Vektoren aufzuspannen, die konjugiert sind, also bezüglich einem Skalarprodukt senkrecht aufeinander stehen. Dieses Skalarprodukt wird dabei durch die Matrix **A** definiert:

$$\langle u, v \rangle_{\mathbf{A}} = u^T \mathbf{A} v \quad (2.3.15)$$

Wenn ein Satz dieser konjugierten Vektoren gefunden ist, kann die Lösung x als Linearkombination dargestellt werden. Die entsprechenden Koeffizienten können aufgrund der Orthogonalität leicht ermittelt werden (Gleichung (2.3.16)). Üblicherweise werden nicht alle Koeffizienten α_i benötigt beziehungsweise berechnet. Man nutzt vielmehr ein iteratives Vorgehen, wobei in jedem Schritt die aktuelle Lösung, das Residuum und die neue Suchrichtung aktualisiert werden.

$$\begin{aligned} x &= \sum_i \alpha_i p_i \Rightarrow \mathbf{A}x = \sum_i \alpha_i \mathbf{A}p_i = b \\ p_k^T b &= \sum_i \alpha_i p_k^T \mathbf{A}p_i = \alpha_k p_k^T \mathbf{A}p_k \\ \alpha_k &= \frac{p_k^T b}{p_k^T \mathbf{A}p_k} \end{aligned} \quad (2.3.16)$$

Die Robustheit der Methode bezüglich Randwerten und die Möglichkeit der Bearbeitung nichtlinearer Probleme erkauft man sich allerdings mit gegenüber den direkten Methoden erhöhtem Rechenaufwand, wie Abbildung 2.18 illustriert. Die Rechenzeit pro Gitterpunkt steigt etwa quadratisch und ist ein bis zwei Größenordnungen über der der direkten Löser. Das größte Problem hierbei stellt die schlechter werdende Konditionierung der Matrix, ein Maß für die Stabilität der Abbildung, mit kleiner werdender Schrittweite dar. Für das vorliegende System steigt die Konditionierungszahl, definiert als Verhältnis von maximalem zu minimalem Eigenwert der Matrix, quadratisch an. Dieses Verhalten resultiert in einem Anstieg der Iterationszahl, was schließlich den starken Anstieg der Rechenzeit bewirkt.

Durch eine sogenannte Vorkonditionierung kann die Stabilität des Lösungsalgorithmus verbessert werden. Dazu multipliziert man die ursprüngliche Gleichung mit einer Konditionierungsmatrix \mathbf{P} , was bei richtiger Wahl zu einer Verbesserung der Konditionszahl des Systems führt:

$$\mathbf{P}(\mathbf{Ax} = \mathbf{b}) \Leftrightarrow \mathbf{PAx} = \mathbf{Pb} \Leftrightarrow \mathbf{Mx} = \mathbf{c} \quad (2.3.17)$$

Falls \mathbf{P} in etwa die Inverse von \mathbf{A} darstellt, ist große Stabilität und schnelle Konvergenz garantiert.

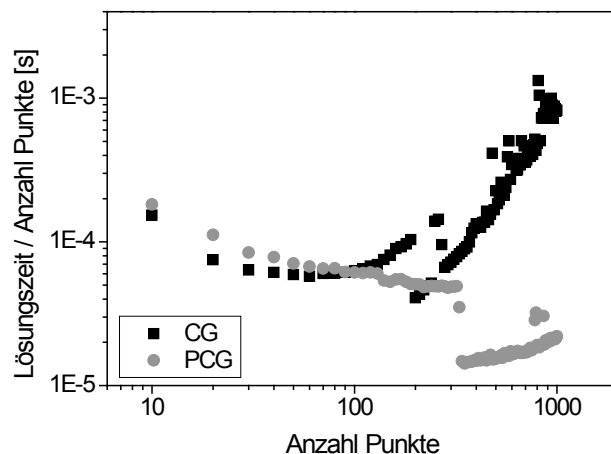


Abbildung 2.18: Lösungszeit pro Gitterpunkt für die Methode der konjugierten Gradienten (CG). Bei Nutzung eines Vorkonditionierers (PCG) steigt der Rechenaufwand nicht.

Wie Abbildung 2.18 illustriert, ist der Rechenaufwand, dargestellt als Rechenzeit pro Stützstelle, bei Nutzung einer Vorkonditionierung für die Methode der konjugierten Gradienten vergleichbar mit dem bei Nutzung direkter Löser (Abbildung 2.17 rechts).

Die iterativen Löser bieten zusätzlich den Vorteil, eine einfache Implementierung nichtlinearer Probleme zu bieten: Anstatt den Fehlervektor in jedem Schritt rekursiv zu berechnen, wird dieser Wert direkt ausgewertet. Da die Algorithmen versuchen, den Fehlervektor zu minimieren, wird nach einigen Iterationen eine zufriedenstellende Näherungslösung erhalten.

Als Beispiel soll wiederum das bereits behandelte Halbleiterproblem ((2.3.5),(2.3.6)) benutzt werden. Bei parametrischer Lösung, analog dem Vorgehen zu Abbildung 2.15, ergeben sich folgende Ergebnisse: Für moderate Gradienten, also Werte für die Änderung der Lösung pro Ortschaftschritt, ist der Fehler der numerischen Lösung auch bei großer Schrittweite gering, was einen geringen Zeitaufwand bedeutet.

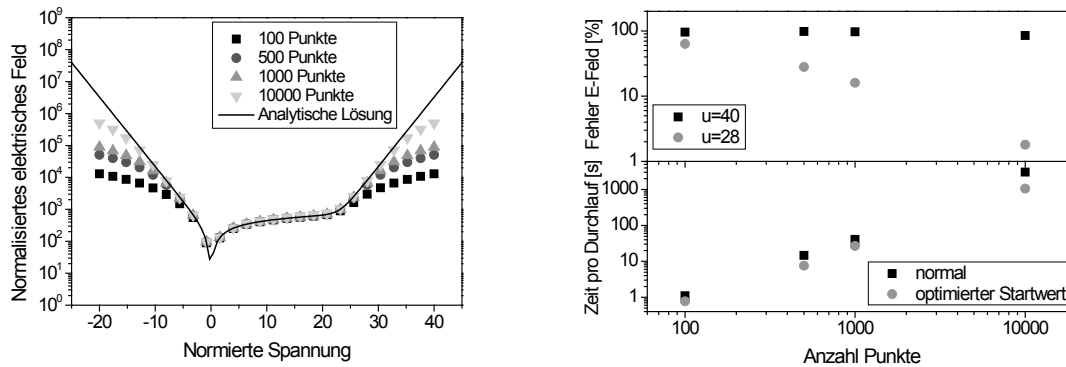


Abbildung 2.19: Parametrische Lösung für das Halbleiterproblem bei Nutzung der Methode der finite Differenzen und einem angepassten PCG-Löser. Die Übereinstimmung der numerischen Lösung mit der analytischen, dargestellt durch das elektrische Feld am linken Rand, ist in einem großen Bereich der Parameter gut (linker Graph). Allerdings weicht die Lösung für große elektrische Felder stark ab, was nur durch eine deutliche Verringerung der Schrittweite vermieden werden kann (rechter Graph).

Dieses in Abbildung 2.19 (linker Graph) dargestellte Verhalten verdeutlicht ein prinzipielles Problem gitterbasierter Methoden: Abhängig von der Schrittweite gibt es einen maximalen Wert für die Ableitung der Lösungsfunktion, der erzeugt werden kann. Gerät man also mit den erwarteten Werten in diesen Bereich, tritt drastische numerische Dämpfung ein, was das Abflachen der dargestellten Kurven erklärt.

Übergang zu Dimensionen grösser als eins

Die bisherigen Ausführungen gelten lediglich für den Fall eindimensionaler Probleme, erfreulicherweise gestaltet sich der Übergang zu höheren Dimensionen relativ einfach. Die 1D-Diskretisierung bezog sich auf einen Satz Indizes $i=1 \dots N$. Im zweidimensionalen Fall hat man, bei äquidistanter Anordnung in beiden Richtungen, allerdings schon ein Paar Indizes pro Punkt, also etwa:

$$x_{ij} = (ih, jh) \quad (2.3.18)$$

Das würde bei simpler Übertragung der Operatoren die Aufgabe von „Löse Matrix mal Vektor gleich Vektor“ zu „Löse Matrix gefaltet mit Matrix gleich Matrix“ umwandeln:

$$h^2 V_{i,j}'' = V_{i-1,j} + V_{i+1,j} + V_{i,j-1} + V_{i,j+1} - 4V_{i,j} \quad (2.3.19)$$

Da sich diese Aufgabenstellung nur schwer behandeln lässt, wird üblicherweise folgender Weg gewählt: Die auftretenden Gitterpunkte werden unnummeriert, wie in Abbildung 2.20 dargestellt. Auf den so neu benannten Gitterpunkten lässt sich wiederum eine Matrix definieren, die den gewünschten Differentialoperator repräsentiert. Im Falle des Laplace-Operators nimmt diese Matrix folgende Form an:

$$h^2 V_k'' = V_{k-1} + V_{k+1} + V_{k-N} + V_{k+N} - 4V_k \quad (2.3.20)$$

Offenbar ist die Matrix wiederum schwach besetzt und banddiagonal, weshalb sich die besprochenen Algorithmen zur Lösung auch nicht entscheidend ändern.

Die Abhängigkeit des Fehlers von der Anzahl der Gitterpunkte pro Länge illustriert Abbildung 2.22. Augenscheinlich skaliert der Fehler mit einem Potenzgesetz, hier etwa invers proportional. Zum Vergleich und zur Zuordnung dieses Fehlers soll hier eine geometriebeschreibende Größe genutzt werden, die Abweichung der diskretisierten Fläche von der Kreisfläche. Auch hier zeigt sich in etwa die gleiche Skalierung mit der Anzahl der Punkte, was nahe legt, den vorliegenden Alias-Effekt als Ursache der Probleme anzunehmen.

Man kann erwarten, und es ist auch aus den präsentierten Daten ersichtlich, dass eine höhere Anzahl von Punkten größere Genauigkeit verspricht. Allerdings muss dieser Vorteil durch immens ansteigenden Speicherbedarf erkaufte werden, wie Abbildung 2.23 zeigt. Der vom Betriebssystem angezeigte vom Rechenprozess belegte Arbeitsspeicher skaliert mit der vierten Potenz der Anzahl der Punkte pro Längenintervall. Obwohl Arbeitsspeicher heutzutage günstig erhältlich ist, stellt das Verfeinern der Geometrie so keinen erstrebenswerten Ausweg dar.

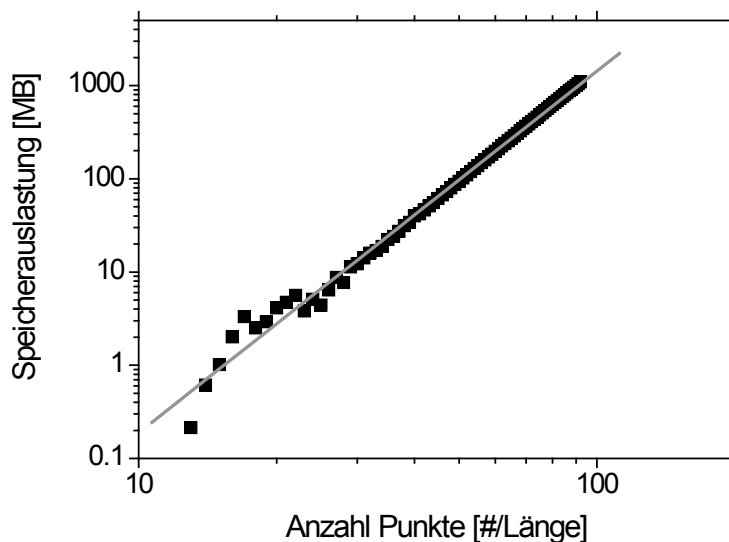


Abbildung 2.23: Abhängigkeit der Speicherbelastung von der Anzahl der Punkte für das Dirichletproblem (2.3.21). Der nichtlinear steigende Speicherbedarf stellt ein ernst zu nehmendes Problem dar - Verfeinerungen des Gitters sollten so nur in Betracht gezogen werden, wenn keine Alternativen verfügbar sind.

Einen eleganten Ausweg aus diesem Dilemma bietet die Methode der finiten Elemente, welche im Folgenden beschrieben werden soll.

2.4 Die Methode der finiten Elemente

Wie in der vorangegangenen Kapiteln gezeigt wurde, ist es mit vertretbarem technischen Aufwand möglich, Randwertprobleme numerisch zu behandeln: In einer Dimension bieten sich hier die Shooting-Methoden an, während in höheren Dimensionen zum Beispiel die Methode der finiten Differenzen genutzt werden kann. Große Nachteile ergeben sich allerdings aus der Definition von in allen Dimensionen äquidistanten Gitterpunkten. Offenbar erscheint es so schwierig, starke Gradienten in den Lösungsfunktionen zu erfassen und krummlinige Ränder abzubilden.

Ausgangspunkt zur Erläuterung der Methode soll wieder das eindimensionale Standardproblem sein:

$$\Delta V = \begin{cases} -1; & x \in (0,1) \\ 0; & \text{sonst} \end{cases}; V(0) = V(1) = 0 \quad (2.4.1)$$

Diese Gleichung kann nun mit einer willkürlichen Funktion⁷ v multipliziert und anschließend über das Problemgebiet Ω integriert werden:

$$\begin{aligned} -\Delta V \cdot v &= v, v \in H_0^1 \\ \int_{\Omega} -\Delta V \cdot v &= \int_{\Omega} v \\ \int_{\Omega} \nabla V \cdot \nabla v &= \int_{\Omega} v \end{aligned} \quad (2.4.2)$$

In (2.4.2) stellt H_0^1 prinzipiell den Raum stetig differenzierbarer Funktion auf $(0,1)$ dar, welche die Randbedingungen erfüllen. Durch die Wahl geeigneter Funktionale wird dieser Raum zu einem Hilbertraum. Bei der Umformung des Integrals wurde die erste Greensche Formel unter Beachtung der verschwindenden Randwerte genutzt. Offenbar wird so ein neues mathematisches Umfeld erzeugt, welches nicht nach Lösungen von (2.4.1) sucht, sondern nach sogenannten schwachen Lösungen, welche über Integrale mit Testfunktionen definiert sind.

Um das Problem nun numerisch lösbar zu gestalten, wird eine geeignete Basis (mit Basisfunktionen v_k) des zugrunde liegenden Funktionenraums gewählt und sowohl die gesuchte Lösung als auch die rechte Seite werden als Linearkombination dieser formuliert:

$$V(x) = \sum_{k=1}^n V_k v_k(x); f(x) = \sum_{k=1}^n f_k v_k(x) \quad (2.4.3)$$

In obiger Gleichung wurde die rechte Seite, welche im vorgestellten Standardproblem konstant ist, durch eine Funktion $f(x)$ ersetzt. Ein Einsetzen dieser Formulierungen in (2.4.2) ergibt:

$$\sum_{k=1}^n V_k \int_{\Omega} \nabla v_k \nabla v_j = \sum_{k=1}^n f_k \int_{\Omega} v_k v_j \quad \forall j \quad (2.4.4)$$

Wie notiert muss (2.4.4) für alle Komponenten k und alle Indizes j gelten. Man kann offensichtlich eine etwa übersichtliche Form dieser Forderung angeben:

$$\begin{aligned} \mathbf{L} \vec{V} &= \mathbf{M} \vec{f} \\ \mathbf{L}_{i,j} &= \int_{\Omega} \nabla v_i \nabla v_j; \mathbf{M}_{i,j} = \int_{\Omega} v_i v_j \end{aligned} \quad (2.4.5)$$

Das Problem reduziert sich somit wiederum auf das Lösen eines Gleichungssystems, Ansätze dazu sind bereits beschrieben worden. Eine geeignete, und durchaus intuitive, Basis wäre für das beschriebene Problem ein Satz stückweise linearer Funktionen. Üblicherweise werden sogenannte Zelfunktionen genutzt.

⁷ Natürlich müssen gewisse Bedingungen an Glattheit und Definitionsbereich gestellt werden.

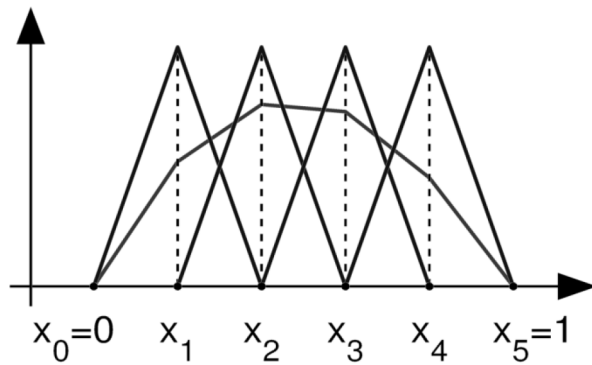


Abbildung 2.24: Zeltfunktionen zur Beschreibung des Vorgehens der finiten Elemente in einer Dimension. Die gewählten Funktionen sind 1 an den jeweiligen Gitterpunkten und verschwinden an den Nachbarknoten.

Nutzt man die in Abbildung 2.24 skizzierten Funktionen auf einem äquidistanten Gitter, werden erstaunlicherweise mathematische Formulierungen erzeugt, welche äquivalent zu den Finiten Differenzen sind. Scheinbar ist somit noch kein Vorteil ersichtlich, aber:

- 1.) Elemente höherer Ordnung können relativ einfach implementiert werden. So können zum Beispiel stückweise quadratische Funktionen genutzt werden, wobei die approximierte Lösung offensichtlich deutlich genauer ist.
- 2.) Komplizierte Gitter sind einfach zu implementieren. Im Gegensatz zu dem recht starren Gerüst der Finiten Differenzen ist ein nicht äquidistantes Gitter wenig problematisch.
- 3.) Eindimensionale Finite Elemente werden prinzipiell nicht genutzt. Die Ausführungen dienen lediglich der Veranschaulichung und offenbaren nicht das gesamte Potential der Methode.

Die Konzepte müssen beim Übergang zu höheren Dimensionen nicht geändert werden, lediglich eine angepasste Aufteilung des Gebietes und nützlich Basisfunktionen gilt es zu finden. Im Standardfall wird das Gebiet in Dreiecke (Abbildung 2.25) zerlegt, auf denen wiederum Zeltfunktionen definiert werden. Diese haben jetzt natürlich eine kompliziertere Form, was, in Kombination der Generierung des Gitters, eine Implementierung schwieriger gestaltet. Aus diesem Grund wird auf die Nutzung verfügbarer Softwarepakete verwiesen.

Abschließend soll noch das Fehlerverhalten der Finiten Elemente Methode am Beispiel des Dirichletproblems für die Einheitsscheibe demonstriert werden. Die Lösung des Problems zeigt wie erwartet eine gewisse Abweichung von der analytischen Lösung, abhängig von der gewählten Anzahl der Punkte. Interessanterweise ist dabei die Skalierung dieses Fehlers im Falle des von „Matlab“ bereitgestellten Löser nicht besser als für Finite Differenzen, ein Verhalten, welches durch die deutlich bessere Abbildung des Randes nicht auftreten sollte. Bei der Nutzung anderer kommerzieller Systeme, hier „Comsol“ tritt dieses Verhalten nicht in Erscheinung, wahrscheinlich geschuldet der Wahl günstigerer Basisfunktionen. Genauere Hintergründe zur FEM-Methode entnimmt man [18].

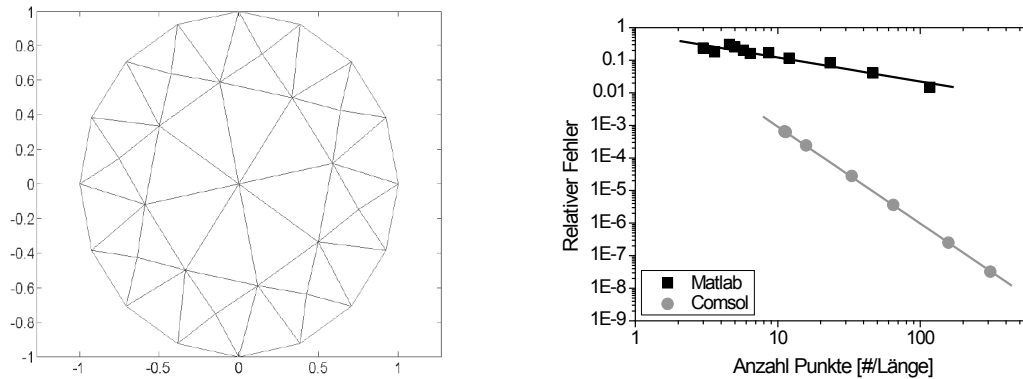


Abbildung 2.25: FEM Gitter für das Dirichletproblem (links) und Entwicklung des Fehlers mit der Anzahl der Gitterpunkte. Offenbar ist es möglich, zumindest bei Wahl bestimmter Software, deutlich bessere Lösungen verglichen mit den finiten Differenzen zu erhalten.

2.5 Numerisches Vorgehen

Sämtliche Berechnungen wurden auf Desktop-Computern durchgeführt, die zum Zeitpunkt der Erstellung dieser Arbeit übliche Leistungsparameter aufwiesen. Beispielsweise wurde dazu ein vierkerniger Prozessor mit etwa 3 GHz Taktrate benutzt, die Größe des Arbeitsspeichers lag bei 4 GB. Die Lösungen wurden teils mittels kommerzieller Software, dem FEM-System „Comsol Multiphysics 3.5“, oder aber mittels selbst implementierter Löser für finite Differenzen ermittelt.

Grundlage aller numerischen Ergebnisse ist die Lösung der Poissongleichung. Unter Anpassung an die tatsächlichen Gegebenheiten in Experimenten muss somit ein Randwertproblem behandelt werden. Die typische Formulierung lautet dabei schematisch:

1. Generation der Geometrie

Abhängig vom Grad der Symmetrie wird das Problem in zweidimensionaler Zylindergeometrie oder dreidimensionaler kartesischer Geometrie modelliert. Das komplette Gebiet kann in drei Bereiche, die Domänen, zerlegt werden:

a) Probe

Im Falle der einfachen Wechselwirkung zwischen Spitze und metallischer Probe entfällt dieser Bereich, da die Wirkung der Oberfläche komplett über Randbedingungen beschrieben werden kann.

Falls zusätzliche Materialien beschrieben werden müssen, zum Beispiel Dielektrika oder Halbleiter, wird in diesen die Ladungsträgerdichte formeltechnisch beschrieben.

b) Raum zwischen Spitze und Probenoberfläche

Dieser Bereich wird als Vakuum betrachtet.

c) Spitze

Abhängig von der Komplexität wird hierbei die Spitze vollständig, also bestehend aus Apex, konischem Anteil und Cantilever, oder nur teilweise modelliert. Falls es aus numerischen Gründen notwendig ist, diesen Bereich als Körper statt als Hülle zu beschreiben, wird eine sehr hohe Dielektrizitätskonstante angesetzt, um Feldfreiheit im Inneren zu generieren.

2. Setzen der Randbedingungen

a) „Unterer“ Rand

Der „untere“ Rand erhält immer eine Dirichletbedingung der Form Potential $U=0$ zugeordnet.

b) Ränder der Spitze

An den Rändern der Spitze gilt eine Dirichletbedingung der Form Potential $U=U_{Tip}$.

c) Übergänge zwischen Materialien

An inneren Rändern gilt Kontinuität.

d) Äußere Begrenzungen der Domäne

An den lateralen äußeren Ränder werden entweder Neumannbedingungen (Symmetrie) oder Dirichletbedingungen $U=0$ gesetzt.

3. Lösen des Problems

Das modellierte Problem wird im Allgemeinen parametrisch gelöst, das heißt, es wird ein Satz von Rechnungen in Abhängigkeit einer der auftretenden Konstanten (zum Beispiel Spitze-Probe-Abstand oder Spannung der Spitze) erzeugt. Für jede Lösung werden dann, je nach Bedarf, folgende Werte erzeugt:

a) Kapazität des Systems

Dazu wird die Flächenladungsdichte, welche proportional zur Normalkomponente des elektrischen Feldes ist, über die Oberfläche der modellierten Spitze integriert. Die so erhaltene Gesamtladung ist proportional zur Kapazität C . Alternativ kann auch die Energiedichte im beschriebenen Gebiet berechnet werden. Beide Möglichkeiten werden in (2.5.1) beschrieben:

$$\begin{aligned} C &= \frac{Q}{U_{Tip}} = \frac{\int_{Spitzenoberfläche} \epsilon_0 E dA}{U_{Tip}} \\ C &= \frac{2W}{U_{Tip}^2} = \frac{\int_{Gebiet} \epsilon_0 E^2 dV}{U_{Tip}^2} \end{aligned} \quad (2.5.1)$$

b) Kraftwirkung auf die Spitze

Dazu wird entweder der Maxwellsche Spannungstensor \mathbf{T} genutzt (siehe auch Gleichung (3.1.2)) oder eine vereinfachte Formulierung, welche die zu Feldstärke und Ladungsdichte proportionale Kraftdichte ausnutzt:

$$\begin{aligned} \vec{F} &= \int_{Spitzenoberfläche} \mathbf{T} d\vec{A} \\ \vec{F} &= \int_{Spitzenoberfläche} \sigma E_{normal} \vec{E} dA \end{aligned} \quad (2.5.2)$$

c) Berechnung spannungsabhängiger „Messgrößen“

Da in tatsächlichen Experimenten, wie beschrieben, Schwingungsamplituden auf verschiedenen harmonischen Frequenzen einer Anregungsfrequenz ω genutzt werden, wird folgender Ansatz verwendet:

$$F_{Tip,z} = F_{Tip,z}(U_{Tip}) = F_{Tip,z}(U_{DC} + U_{AC} \sin(\omega t))$$

$$F_{Tip,z} \approx F_{Tip,z}(U_{DC}) + \frac{\partial F_{Tip,z}}{\partial U_{Tip}} U_{AC} \sin(\omega t) + \frac{1}{2} \frac{\partial^2 F_{Tip,z}}{\partial U_{Tip}^2} U_{AC}^2 \sin^2(\omega t) + \dots \quad (2.5.3)$$

Offenbar genügt es also zur Verknüpfung simulierter Werte mit der Realität, die entsprechenden Ableitungen zu bestimmen⁸. Diese Ableitungen, beziehungsweise Näherungen dafür, werden numerisch durch Bildung des zentralen Differenzenquotienten bestimmt:

$$\frac{\partial F_{Tip,z}(U_i)}{\partial U_{Tip}} \approx \frac{F(U_{i+1}) - F(U_{i-1}))}{U_{i+1} - U_{i-1}} \quad (2.5.4)$$

In (2.5.4) ist i der Zählindex der generierten Kraft- und Spannungswerte. Höhere Ableitungen können durch sukzessive Anwendung gewonnen werden. Der zentrale Differenzenquotient wird hierbei genutzt, da eine bessere Approximationsordnung als bei Verwendung einseitiger Differenzen vorliegt. Es stellt sich heraus, dass die so gewonnenen Ableitungen quadratischer Interpolation entsprechen.

Um einen zu SKM-Experimenten vergleichbaren Wert zu erzeugen, muss das Minimum vom Betrag der ersten Ableitung gefunden werden, eine punktweise Auswertung kann allerdings bei komplexen Kraft-Spannungs-Kurven zu Abweichungen im Bereich der Schrittweite führen. Um das zu umgehen, wird entweder die gesamte Kurve quadratisch angepasst (für Metalle) und das Kraftminimum analytisch aus den Koeffizienten bestimmt oder aber die generierte Kurve wird durch Interpolationswerte ergänzt (für Halbleiter).

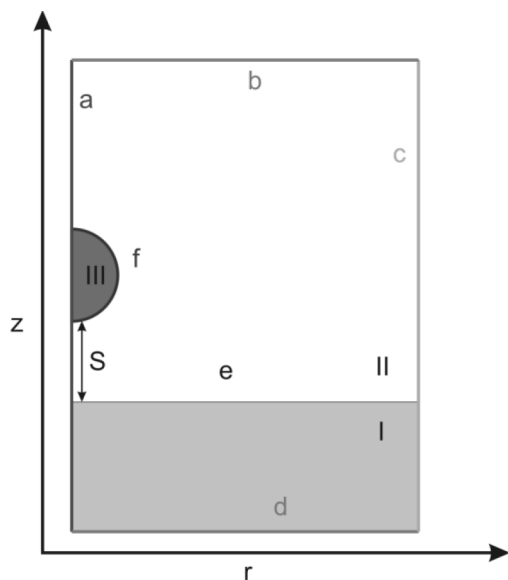


Abbildung 2.26: Schema des Simulationssetups. Die für die Gebiete (römische Ziffer) und Ränder (Buchstaben) benutzten Bedingungen entnimmt man Tabelle 1.

Bezeichner	Bedingung
I	$r=sc(U)$
II	$r=0$
III	$r=0$
a	Axialsymmetrie
b	$U=0$
c	Symmetrie
d	$U=0$
e	Kontinuität
f	$U=U_{Tip}$

Tabelle 1: Benutzte Bedingungen für die Simulation.

⁸ Tatsächlich gilt nur Gleichheit für quasistatische Systeme, also solche realen Systeme, bei denen die Reaktionszeiten des elektronischen Systems deutlich kürzer sind als die Periodendauern der anregenden Schwingung. Für Metalle ist das üblicherweise erfüllt, für Halbleiter leider nur bedingt.

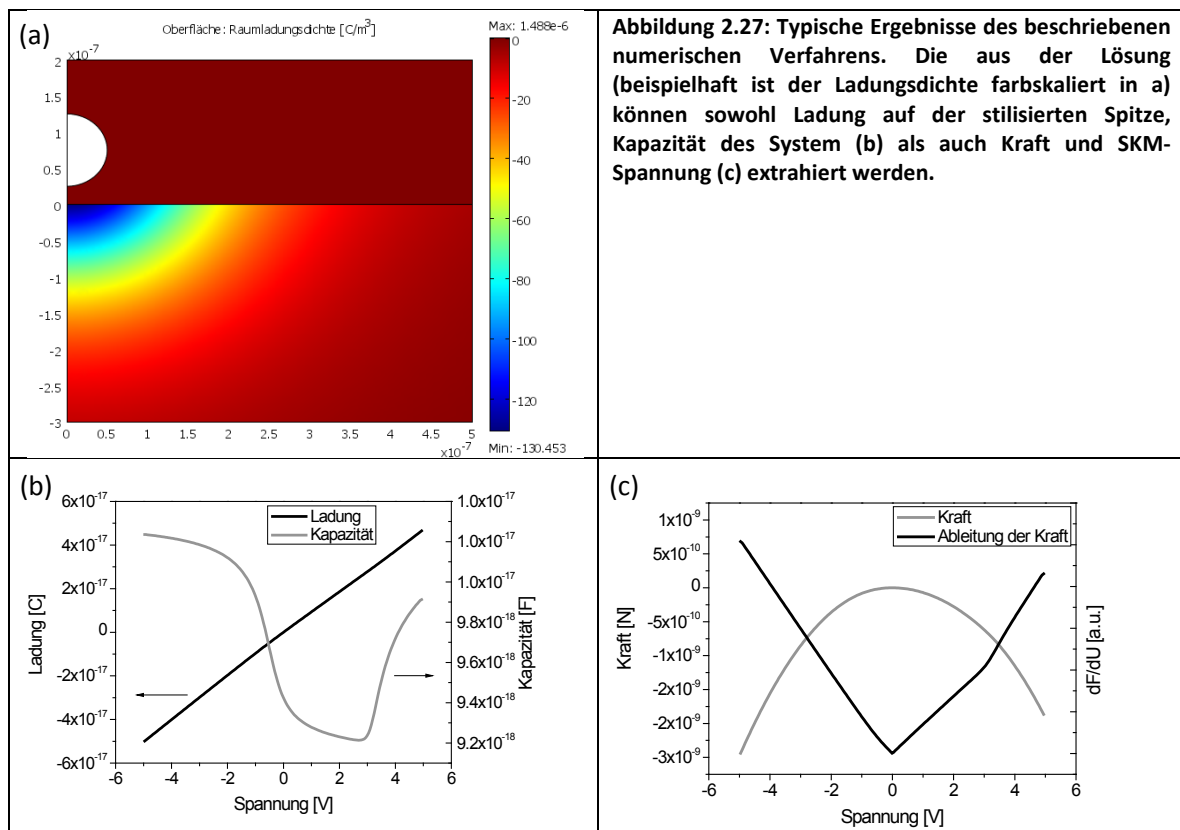
Konzeptionelle Grundlagen

Illustriert werden soll der zuvor beschriebene Ablauf für den Fall „metallische Kugel über Halbleiter“. Offenbar genügt es, Zylindersymmetrie zu nutzen. Eine Skizze des modellierten Setups zeigt Abbildung 2.26. Die zugrunde liegenden Differentialgleichungen sind in II und III die Laplacegleichung, sowie in I die Poissongleichung mit folgender Ladungsdichte sc (von engl. „space charge“):

$$\rho = sc(U) = -2qn_i \left(\sinh\left(\frac{U - U_0}{U_T}\right) + \sinh\left(\frac{U_0}{U_T}\right) \right) \quad (2.5.5)$$

In (2.5.5) stehen ρ für die Ladungsdichte, q für die Elementarladung, n_i für die intrinsische Ladungsträgerdichte in Silizium, U_0 für ein Referenzpotential und U_T für das Spannungsäquivalent der thermischen Energie (für eine genauere Beschreibung dieser Gleichung siehe Anhang A).

Die typischen Ergebnisse sind in Abbildung 2.27 dargestellt. Der Verlauf der aus der Ladung berechneten Kapazität zeigt den erwarteten Verlauf für eine MOS-Struktur, der Verlauf der Kraft-Spannungs-Kurve spiegelt deutlich die Spannungsabhängigkeit der Ladungsträgerdichte wieder: Für betragsgroße Potentiale auf der Spitzenoberfläche verhält sich die Probenoberfläche nahezu metallisch, was wiederum eine rein quadratische Abhängigkeit der Kraft erzeugt.



2.6 Numerische Fehler

Da nicht zu erwarten ist, dass die numerischen Methoden exakte Lösungen liefern, sollen hier einige typische Fehler besprochen werden. Es geht dabei nicht um Abschneidefehler, wie sie bei der Auswertung von zum Beispiel Ableitungen entstehen. Vielmehr sollen Auswirkungen der Geometrie auf die Qualität der Lösung behandelt werden.

Zunächst soll dazu ein sehr einfaches Testproblem definiert werden: Löse die Laplacegleichung in zwei Dimensionen mit einer stufenförmigen Potentialverteilung der Breite b in einem Rechteck der Breite $2W$ und Höhe W . Zur Bewertung wird das errechnete Potential an einem definierten Punkt (hier willkürlich (0,1)) ausgewertet und mit dem analytischen Wert verglichen. Die Symmetrie des Problems wird bei der Berechnung dabei ausgenutzt, demnach muss nur die Lösung in einem Quadrat berechnet werden. In Abbildung 2.28 sind eine typische Lösung und die Entwicklung des relativen Fehlers mit den Parametern b und W angegeben.

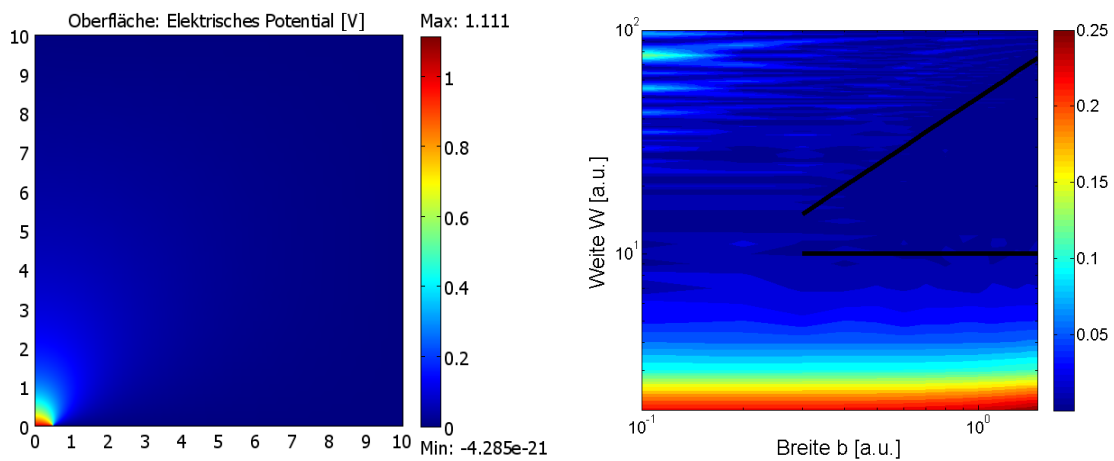


Abbildung 2.28: Typische Lösung des Testproblems (links) und Entwicklung des relativen Fehlers (rechts). Offenbar gibt es einen Bereich „guter“ Werte, in dem der Fehler sehr gering ist. Dieser Bereich ist im Fehlerbild durch schwarze Linien begrenzt.

Der Fehler wurde durch Vergleich mit der analytischen, durch Anwendung der Integralformel (siehe Kapitel 4.2.2) erhältlichen, Lösung am Punkt (0,1) bestimmt:

$$Fehler = \frac{V(0,1) - \frac{2}{\pi} \arctan\left(\frac{b}{2}\right)}{\frac{2}{\pi} \arctan\left(\frac{b}{2}\right)} \quad (2.6.1)$$

Offenbar scheint es drei Effekte zu geben:

1. Unabhängig von der Breite b sinkt der Fehler nichtlinear mit der Weite W (Abbildung 2.29 links). Dieses Verhalten liegt darin begründet, dass die angewendete Integralformel nur für unendliche Ausdehnung gilt. Bei endlicher Ausdehnung gibt es zusätzlich einen Einfluss des oberen Randes, was prinzipiell zu einer exponentiell gedämpften Abweichung führt. Verdeutlicht ist das in Abbildung 2.28 (rechts) durch die willkürlich platzierte Linie $W=10$. Augenscheinlich ist der Fehler für Weiten grösser als diese Schwelle gering, was auch bei Skalierung gelten sollte. Es kann also davon ausgegangen

werden, dass die Ausdehnungen der numerischen Domäne im Bereich von zehnmal der Größe des Meßbereichs, hier eins, liegen sollte.

2. Für Weiten grösser als die so definierte Schwellweite gibt es eine minimale Breite, für welche der Fehler absinkt (Abbildung 2.29 Mitte). In Abbildung 2.28 (rechts) ist dies ebenfalls durch eine Linie der Form $W=50b$ verdeutlicht. Dieses Verhalten ist in fehlender Abtastgenauigkeit begründet: Der Löser „sieht“ die Potentialverteilung nur dann, wenn Gitterpunkte innerhalb der halben Breite b liegen. Somit dürfen die Ausdehnungen der numerischen Domäne maximal etwa fünfzigfach grösser als Breiten von Störungen sein, um diese noch als solche abbilden zu können. Allerdings ist das eher als Vorschrift für die Dichte der Gitterpunkte zu verstehen. Bei den Standardeinstellungen der verwendeten Software sind so etwa fünf Punkte im Bereich der Störung nötig.
3. Der Fehler oszilliert bei konstanter Breite mit der Weite (Abbildung 2.29 rechts). Dieser Effekt kann durch ein ungünstiges Verhältnis von Abtastwellenlänge zu vorhandenen Spektralkomponenten der Störung verstanden werden und kann leider kaum umgangen werden.

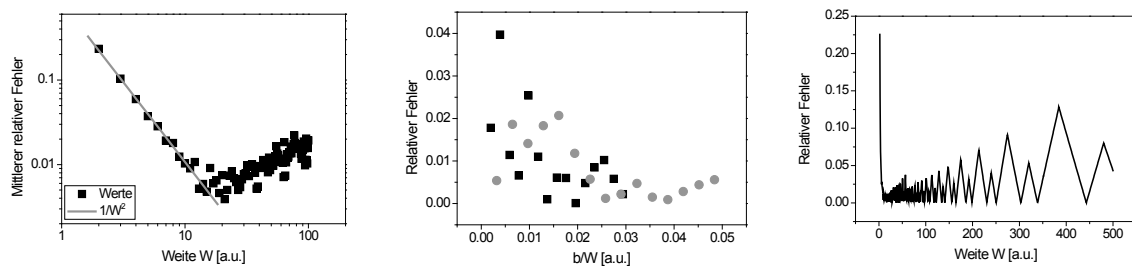


Abbildung 2.29: Typische Skalierungen des relativen Fehlers mit den geometrischen Parametern. Offenbar müssen bestimmte Schwellverhältnisse, sowohl in Weite W der Domäne als auch im Verhältnis der Strukturbreite b zur Weite, beachtet werden, um den Fehler gering zu halten.

Ein weiteres, sehr anschauliches und auch für die im Rahmen dieser Arbeit ermittelten Werte wichtiges, Testproblem ist die Kapazitätsbestimmung einer Kugel. Dieses Problem wurde in Zylindersymmetrie behandelt, wobei die Ladung der Kugel über Integration der Ladungsdichte auf der Oberfläche erhalten wurde. Zur Illustration ist hier der Fehler in Bezug auf die korrekte Formulierung (Gleichung 4.1.9) in Abhängigkeit von der Größe der Domäne gewählt.

Wie in Abbildung 2.30 (rechts) veranschaulicht, fällt auch für dieses Problem der relative Fehler der numerisch berechneten Kapazität mit der Weite der Domäne. Offenbar zwingt das Einbringen des geometrischen Objekts den Löser, eine genügend hohe Gitterdichte zu nutzen, weshalb Abtastfehler nicht auftreten. Die Schwellweite scheint ebenfalls bei etwa zehn Radien der modellierten Kugel zu liegen, was die zuvor gemachten Aussagen unterstützt.

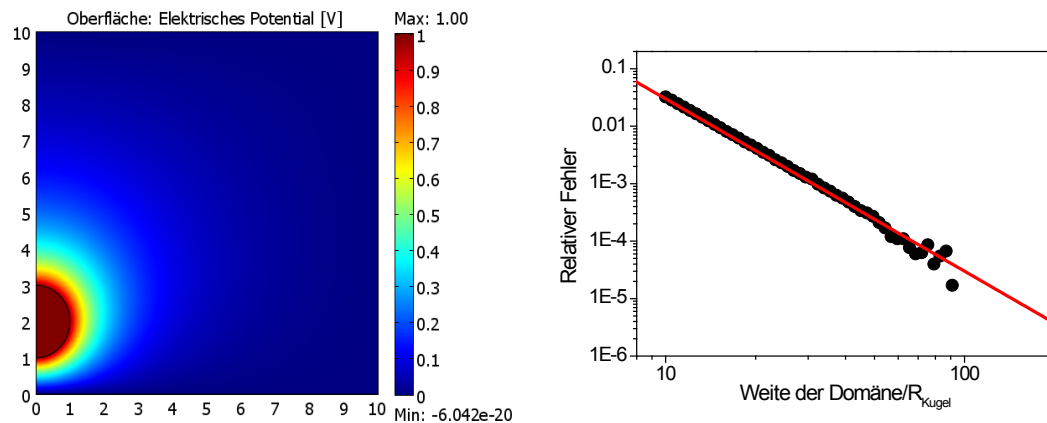


Abbildung 2.30: Typische Lösung des Kugelproblems (links) und Skalierung des relativen Fehlers in der Kapazität mit der Weite der Domäne. Auch in diesem Fall ergibt sich ein nichtlineares Absinken des Fehlers.

Es ist also davon auszugehen, dass der numerische Fehler klein ist, wenn:

1. Die Ränder weit genug entfernt von den interessanten Objekten liegen; ein Verhältnis von etwa zehn zwischen Abstand und Größe sollte eingehalten werden.
2. Randwertfunktionen genügend aufgelöst werden.

Beide Bedingungen sind relativ schwierig gleichzeitig zu erfüllen, da bei der Modellierung oft sehr unterschiedliche Größenordnungen auftreten.

3. Grundlegende Wechselwirkungen im System Spitze-Probe

In diesem Kapitel sollen grundlegende Phänomene im System Spitze-Probe untersucht werden. Von größter Bedeutung erscheint dabei die Untersuchung der Parameterabhängigkeit der wichtigen Messgrößen Kraft und Kapazität. Dazu werden typische Experimente numerisch nachgestellt und aus den Lösungen die zu erwartenden Größen extrahiert.

3.1 Generelle Abhängigkeiten im System Spitze-Metall

In diesem einfachen System können zunächst grundlegende Untersuchungen vorgenommen werden. Dazu wird vereinfachend Zylindersymmetrie angenommen, da für die vorgenommenen Rechnungen keine lateralen Asymmetrien vorlagen. Die Spitze wurde bestehend aus kugelförmigem Apex (Radius R), kegelförmigen Teil (Höhe H , Öffnungswinkel δ) und Basis (Radius R_B) modelliert. Der Übergang zwischen Apex und Kegel wurde dabei so gewählt, dass die Randkurve glatt ist. Es wurden daher eher abgerundete Kegel untersucht (bei den üblicherweise kleinen Öffnungswinkeln besteht der Apex fast aus einer Halbkugel). Abbildung 3.1 zeigt die verwendete Form.

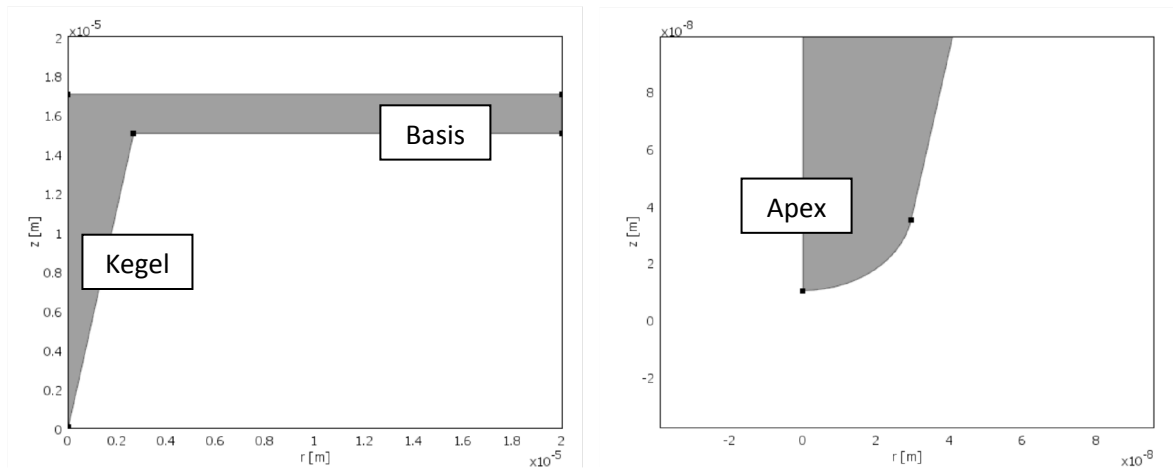


Abbildung 3.1: In den Simulationen genutzte Form der Spitze (hier $R=30$ nm, $H=15$ μ m, $\delta=10^\circ$, $R_B=20\mu$ m). Im rechten Bild ist der Bereich des Apex vergrößert dargestellt, um den gewählten Verlauf des Übergangs zum Kegel zu verdeutlichen.

Die Randbedingungen wurden wie folgt gesetzt: Erdung für $z=0$, Axialsymmetrie um $r=0$, Spannung auf den Rändern der Spitze konstant $=V_T$, Erdung für die verbleibenden Ränder. Die Berechnungen wurden parametrisch durchgeführt, das heißt, die feste Spannung V_T auf den Rändern der Spitze wurde variiert.

Die Ladung, sowohl der gesamten Konstruktion als auch von den Teilbereichen, wurde dann durch Integration der Normalkomponente der dielektrischen Verschiebung D auf den Rändern bestimmt:

$$Q = \int_A \sigma dA = \int_A D_{normal} dA = \int_A \epsilon_0 E_{normal} dA \quad (3.1.1)$$

Durch lineare Anpassung der Ladung an die Spannung der Spitze wurden die Kapazitäten bestimmt, da für dieses System $Q=CV_T$ gelten muss.

Grundlegende Wechselwirkungen im System Spitze-Probe

Für die Berechnung der Kräfte wurde ähnlich verfahren, allerdings musste hierbei eine umfangreichere Rechnung durchgeführt werden, da gilt:

$$\vec{F} = \int_A \mathbf{T} d\vec{A} \quad (3.1.2)$$

$$\mathbf{T}_{ij} = \varepsilon_0 E_i E_j + \frac{1}{\mu_0} B_i B_j - \frac{1}{2} \left(\varepsilon_0 E^2 + \frac{1}{\mu_0} B^2 \right) \delta_{ij}$$

In (3.1.2) steht \mathbf{T} für den Maxwell'schen Spannungstensor, der die elektromagnetische Kraftwirkung auf Oberflächen beschreibt [19]. Allerdings wurde der magnetische Anteil, gegeben durch die Komponenten der magnetischen Flussdichte B , nicht beachtet. In den durchgeführten statischen Berechnungen tritt dieser ohnehin nicht auf, tatsächlich sollte dieser Anteil auch in dynamischen Messungen mit elektrischen Wechselfeldern keine Rolle spielen, wie folgende Abschätzung zeigt:

$$\oint B dr = \frac{1}{c^2} \int \frac{\partial E}{\partial t} dA \quad (3.1.3)$$

In Gleichung (3.1.3) wurde der Stokesche Satz auf die Maxwellgleichung angewendet, welche Flussdichte und elektrisches Feld verknüpft. Wird vereinfachend eine Halbkugel als Oberfläche gewählt, ergibt sich weiter:

$$2\pi R B = 2\pi R^2 \frac{1}{c^2} \frac{\partial E}{\partial t} \quad (3.1.4)$$

$$\hat{B} = \frac{1}{2} \frac{1}{c^2} \omega \hat{E} R$$

Die so eingeführten, mit „Dach“ gekennzeichneten Größen, sind die Amplituden der entsprechenden Felder, ω entspricht der Kreisfrequenz der angelegten Spannung. Berechnet man nun weiter das Verhältnis der Einträge in (3.1.2), erhält man:

$$\frac{\frac{1}{\mu_0} B^2}{\varepsilon_0 E^2} = \frac{1}{4c^2} R^2 \omega^2 \approx \frac{(50 \cdot 10^{-9} \text{ m})^2 (150 \cdot 10^3 \text{ s}^{-1})^2}{4 \cdot (3 \cdot 10^8 \text{ ms}^{-1})^2} \approx 10^{-22} \quad (3.1.5)$$

Dabei wurden typische Größenordnungen für Frequenzen und Dimensionen angenommen. Das erhaltene Verhältnis ist sehr klein, selbst bei Skalierung der Integrationsflächen sollte also kein nennenswerter Beitrag der magnetischen Kraftkomponente zu erwarten sein.

Die erhaltenen Kräfte wurden quadratisch an die Spannung angepasst. Die extrahierten Koeffizienten für den quadratischen Term sollten gut zur Beschreibung der Kraftwirkung nutzbar sein, da die wirkende Kraft als Produkt vom Quadrat der Spannung und einem Geometrieterm darstellbar ist:

$$F = k(s, R, H, \dots) \cdot V_T^2 \quad (3.1.6)$$

Formulierung (3.1.6) gewinnt man zum Beispiel durch Betrachtung von (3.1.2). So kann man das elektrische Feld als $E=V_T/l$ definieren, wobei l eine, vom jeweiligen Ort abhängige, Länge der Feldlinie bedeutet.

Die Ergebnisse für eine Studie der Abhängigkeit von der Separation S zwischen Spitze und Probe zeigt Abbildung 3.2. Dabei wurden typische Parameter für die entsprechenden geometrischen Größen genutzt.

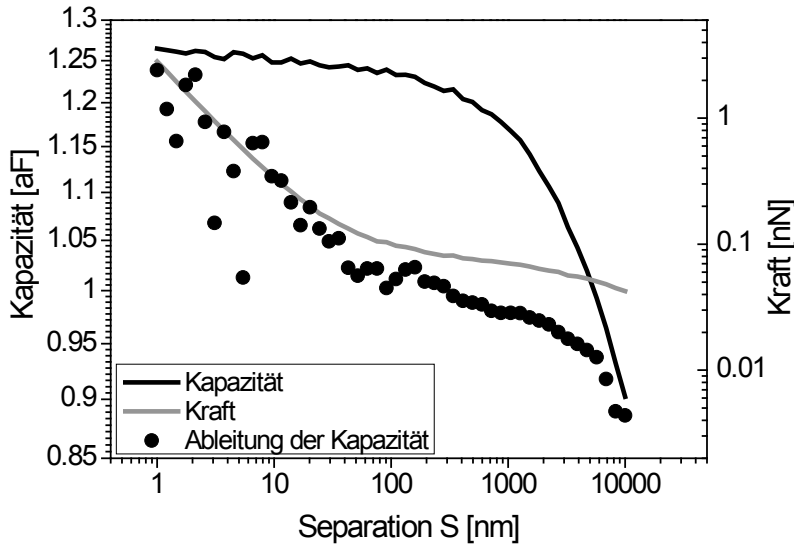


Abbildung 3.2: Extrahierte Kapazitäten und Kräfte auf eine realistisch modellierte Spitze. Offenbar ist die Kapazität des Systems für moderate Separationen quasi konstant, während die Kraft eine deutliche Abhängigkeit zeigt. Des Weiteren liegt gute Übereinstimmung zwischen Ableitung der Kapazität und Kraft vor, allerdings durch numerische Artefakte verschleiert.

Offensichtlich zeigen sowohl Kapazität und als auch Kraft eine gewisse Abhängigkeit von der Separation, die Verläufe sind dabei allerdings nicht vergleichbar: Während die Kapazität fast konstant erscheint, ändert sich die Kraft über zwei Größenordnungen. Zusätzlich dargestellt ist der Wert der Ableitung der Kapazität nach der Separation dC/dS , welcher augenscheinlich identisch mit der Kraft ist. Erklärbar ist dieses Ergebnis, wenn man folgendes Gedankenexperiment durchführt: Bei Verschiebung der Spitze um einen kleinen Wert ΔS ändert sich die Energie W des Systems gemäß:

$$W(S + \Delta S) = \frac{1}{2} C(S + \Delta S) V_T^2 \quad (3.1.7)$$

Fasst man nun die Energieänderung als verrichtete Arbeit auf, ergibt sich:

$$\begin{aligned} F \Delta S &= W(S + \Delta S) - W(S) = \frac{1}{2} C(S + \Delta S) V_T^2 - \frac{1}{2} C(S) V_T^2 \\ F &= \frac{1}{2} V_T^2 \frac{C(S + \Delta S) - C(S)}{\Delta S} \approx \frac{1}{2} V_T^2 \frac{dC}{dS} \end{aligned} \quad (3.1.8)$$

Vergleicht man die Anteile von Apex, Kegel und Basis an Kapazität und Kraft, werden die gefundenen Unterschiede erklärbar (Abbildung 3.3). Offenbar wird die Kapazität von der Basis dominiert, bei der Kraft ist das nicht der Fall. Der Grund für dieses Verhalten ist die stark unterschiedliche Abstandsabhängigkeit der Kapazitäten, folgende Näherungen können genutzt werden (siehe Kapitel 4.1):

$$\begin{aligned} C_{Apex} &\sim R_{Apex} \cdot \ln\left(1 + \frac{R_{Apex}}{S}\right) \\ C_{Kegel} &\sim S \cdot \ln\left(\frac{S}{S+H}\right) \\ C_{Basis} &\sim \frac{1}{H+S} \end{aligned} \quad (3.1.9)$$

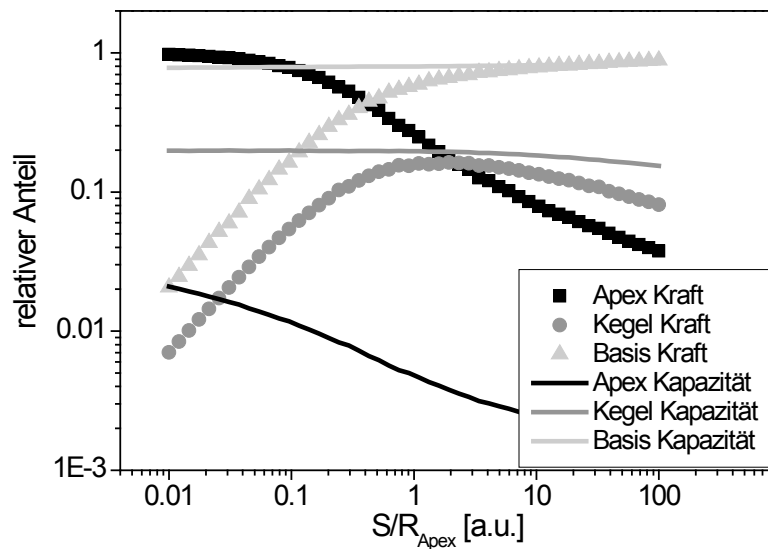


Abbildung 3.3: Relative Anteile von Apex, Kegel und Basis an Kapazität und Kraft. Die quasi konstante Kapazität der Basis dominiert die Gesamtkapazität, während für kleine Separationen der Apex die Kraft bestimmt.

Eine Darstellung der Verläufe von Funktionen gemäß (3.1.9) sind zusammen mit den aus der Simulation gewonnenen Daten in Abbildung 3.4 dargestellt. Die Übereinstimmung ist mehr als zufriedenstellend, was die Nutzbarkeit der Formulierungen deutlich unterstreicht. Setzt man zur Abschätzung typische geometrische Dimensionen an, ergibt sich so für die Verhältnisse:

$$\begin{aligned} \frac{C_{Apex}}{C} &\approx \frac{R_{Apex}}{H} \ln\left(1 + \frac{R_{Apex}}{S}\right) \\ \frac{C_{Kegel}}{C} &\approx \frac{S}{H} \ln\left(\frac{S}{H}\right) \end{aligned} \quad (3.1.10)$$

Repräsentative Zahlenwerte sind demnach für den Apex in der Größenordnung von 1%, für den Kegel ergeben sich Werte von etwa einigen Prozent, wie man auch Abbildung 3.3 entnehmen kann. Bereits an dieser Stelle lässt sich ein, vorsichtiges, Zwischenfazit ziehen: Ausgehend von den Beschreibungen scheint die Spitze eines AFM-Systems nicht allzu gut für

Kapazitätsmessungen geeignet, da die relativ große Basis einen sehr großen Anteil liefert, was im Endeffekt große räumliche Mittelung bedeutet.

Im Gegensatz dazu verspricht die Kraftmessung höhere räumlich Auflösung, da sie am Apex konzentriert scheint. Ebenfalls in Kapitel 4.1 werden Näherungen für die Kraft gegeben:

$$\begin{aligned} F_{Apex} &\sim R_{Apex} \left(\frac{1}{S} - \frac{1}{S + 3R_{Apex}} \right) \\ F_{Kegel} &\sim \ln \left(\frac{H}{S + 3R_{Apex}} \right) \\ F_{Basis} &\sim \left(\frac{R_{Basis}^2}{H^2} - 1 \right) \end{aligned} \quad (3.1.11)$$

Durch Vergleich von (3.1.9) mit (3.1.11) wird das unterschiedliche Verhalten von Kapazität und Kraft ersichtlich. Interessant ist auch hier eine Abschätzung der Anteile an der Gesamtkraft, ebenfalls für übliche Größen. So ergeben sich hierbei für den Apexanteil für kleine Separationen im Bereich des Apexradius Werte von etwa 100%, was durchaus den Gedanken an hohe Auflösung motiviert.

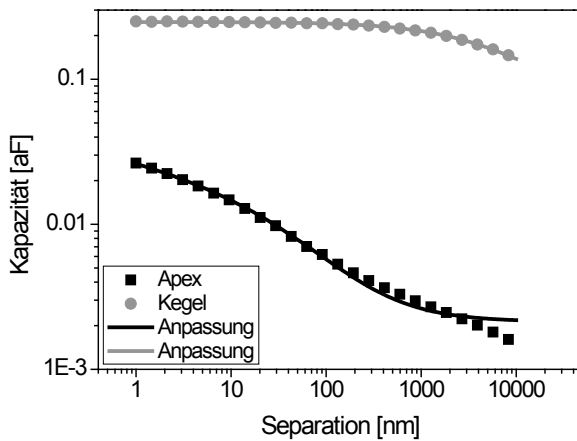


Abbildung 3.4: Vergleich der Näherungsfunktionen für Apex- und Kegelkapazität mit den numerischen Ergebnissen. Die Übereinstimmung ist über einen weiten Bereich von Separationen gut.

Es erscheint nun günstig, die Skalierung der Ergebnisse mit bestimmenden Parametern zu betrachten. In diesem Fall sind das die Abhängigkeiten von Apexradius, Höhe des Kegels und Weite der Basis.

Abbildung 3.5 zeigt Ergebnisse der Untersuchung auf Abhängigkeit vom Apexradius. Die Kapazität des Systems wird maßgeblich von der Basis bestimmt, weshalb sich Änderungen dieses Radius kaum auswirken.

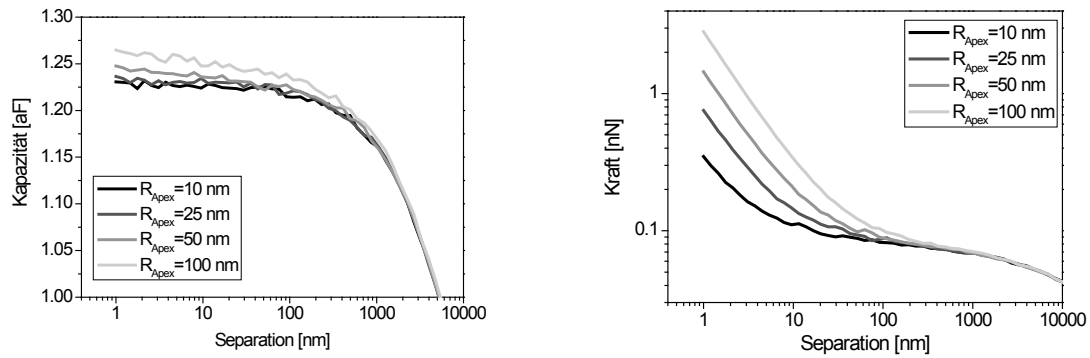


Abbildung 3.5: Abhängigkeit der Kapazität (links) und der Kraft (rechts) vom Radius des Apex, R_{Apex} . Während sich die Kapazitätskurven nur minimal ändern, ergibt sich für die Kraft, in gewissen Bereichen für die Separation, ein deutlicher Unterschied.

Ein deutlich anderes Bild ergibt sich bei der Kraft: Die Verläufe der Kraft-Abstands-Kurven skalieren deutlich mit dem Radius, was aus den zuvor besprochenen Näherungen ableitbar ist. Interessant ist eine Darstellung über der normierten Separation, definiert als Verhältnis von Separation und jeweiligem Radius (Abbildung 3.6). Alle dargestellten Kraft-Abstandskurven liegen in diesem Plot auf einer universellen Kurve, welche die einfache Abhängigkeit $F \sim \text{const} + (s^2 + s)^{-1}$ hat. Das kann sofort aus (3.1.11) gefolgert werden und muss hier nicht weiter erläutert werden.

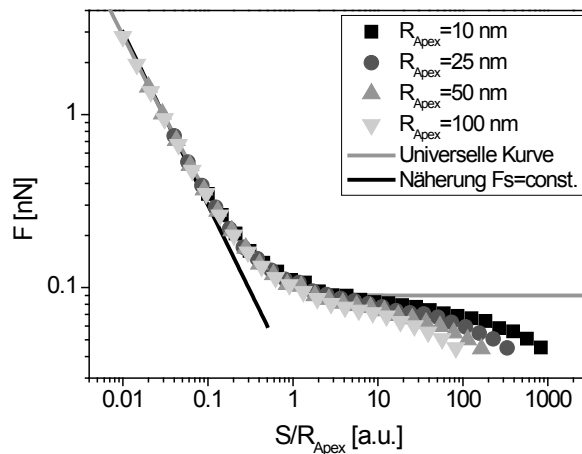


Abbildung 3.6: Kraft-Abstands-Kurven in normierter Darstellung. Alle individuellen Kurven (Symbole) liegen auf einer universellen Linie (grau), die die prinzipielle inverse Proportionalität der Kraft zur Separation darstellt. Die in grau gezeichnete Kurve verdeutlicht diesen Zusammenhang.

Da, wie bereits mehrfach erwähnt, die Gesamtkapazität des Systems fast ausschließlich von der Fläche der Basis bestimmt wird, diese aber im Gegensatz wenig zur Kraft beiträgt, sollen im Weiteren nur noch die Auswirkungen der Parameter des kegelförmigen Teils auf die Kraft besprochen werden. Wie Abbildung 3.7 zu entnehmen ist, ändern Variationen der Parameter Höhe und Öffnungswinkel die Ergebnisse kaum. Da ein Großteil kommerziell verfügbarer Cantilever sehr ähnliche Werte dieser Parameter besitzt, kann im Rahmen einer Modellierung durchaus ein mittlerer Wert genutzt werden.

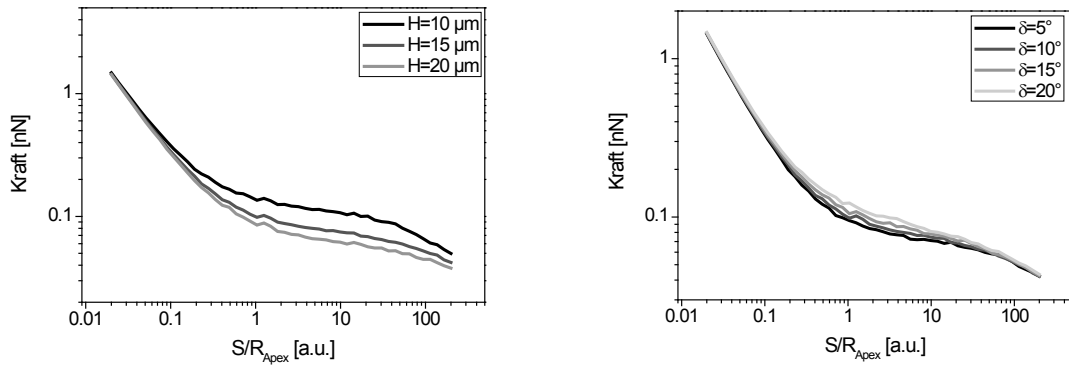


Abbildung 3.7: Einfluss der Kegelparameter auf die Kraft. Beide maßgeblichen Parameter, Höhe H (links) und halber Öffnungswinkel δ (rechts) erzeugen nur geringe Änderungen der Kraft-Abstandskurven. Für Werte in üblichen, durch Hersteller abgedeckten, Bereichen und für in Experimenten benutzte Separationen kann dieser Einfluss also durchaus als Konstante in die Kraft einbezogen werden.

Des Weiteren erlaubt es die Form der Kraft-Abstands-Kurven, den Einfluss des Kegels als Konstante in der Kraft zu beschreiben, was eine Entwicklung von nützlichen Abschätzungen unterstützt.

Aus den bis hierher vorgestellten Ergebnissen kann somit gefolgert werden:

1. Die Kraft ist deutlich sensibler gegenüber Abstandsänderungen im System als die Kapazität.
2. Für Separationen kleiner als der Radius des Apex folgt die Kraft-Abstands-Abhängigkeit einer universellen Kurve.
3. Der Einfluss der Parameter des kegelförmigen Teils ist gering und kann als Konstante in die Kraft eingefügt werden.

Abschließend sollen an dieser Stelle noch veranschaulichende Parameter eingeführt werden, die möglicherweise die Beschreibung auftretender Phänomene erleichtern. Ausgehend von den bekannten Formel für kreisförmige Plattenkondensatoren gilt:

$$\begin{aligned} \frac{C}{\left| \frac{\partial C}{\partial d} \right|} &= \frac{\varepsilon \frac{A}{d}}{\varepsilon \frac{A}{d^2}} = d \\ C \frac{d}{\varepsilon} &= A = \pi r^2 \end{aligned} \quad (3.1.12)$$

Mit Hilfe der Formulierungen aus (3.1.12) lassen sich effektive Separationen S_{eff} und Radien R_{eff} (Flächen) definieren:

$$\begin{aligned} S_{eff} &= \frac{C}{\left| \frac{\partial C}{\partial S} \right|} = \frac{C}{\left| 2F(S) \right|} \\ R_{eff} &= \sqrt{\frac{CS_{eff}}{\varepsilon \pi}} \end{aligned} \quad (3.1.13)$$

Eine Darstellung dieser Werte, gewonnen aus den Daten, die für Abbildung 3.2 genutzt wurden, zeigt Abbildung 3.8. Offenbar steigen sowohl effektive Separation und effektive Fläche proportional zur Separation. Mit Hilfe der in (3.1.13) definierten Werte lassen sich einfache Abschätzungen, zum Beispiel der flächenbezogenen Kapazität, erstellen oder erste, vorsichtige, Formulierungen über das Auflösungsvermögen machen.

So kann man ausgehend von den Näherungen das Auflösungsvermögen grob durch den effektiven Radius definieren, allerdings entnimmt man Abbildung 3.8 selbst für Spitze-Probe-Abstände in der Größenordnung von 10% des Apexradius Werte von etwa 100 Apexradien. Dies widerspricht doch deutlich den experimentellen Ergebnissen, welche Auflösung im Bereich kleiner Vielfacher des Radius vermuten lassen.

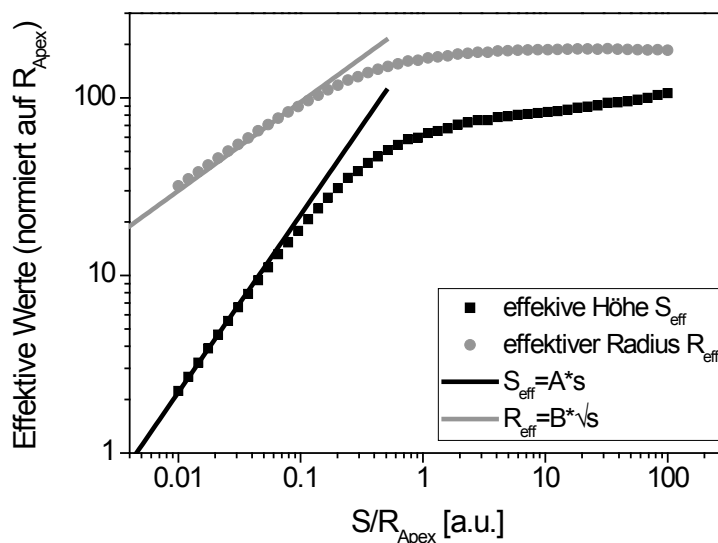


Abbildung 3.8: Effektive Werte für Separation und Radius gemäß (3.1.13). Für kleine relative Separationen lassen sich einfache Zusammenhänge angeben, welche im Endeffekt einfache Abschätzungen erlauben.

3.2 Einfluss von dielektrischen Schichten auf die Messwerte

Von großem experimentellen Interesse ist die Frage nach dem Einfluss möglicherweise geladener dielektrischer Schichten. Die aus einer Messung zu extrahierenden Parameter wären dabei zum Beispiel die fixierte Ladungsträgerdichte und die Dielektrizitätskonstante der Schicht. Beide Werte sind zugänglich, Genauigkeit und prinzipielle Eigenschaften der Abbildung sollen in diesem Abschnitt besprochen werden.

Zur Modellierung wurde dazu zwischen der unteren Begrenzung des Systems und der Spitze eine zusätzliche Schicht der Dicke d mit der relativen Dielektrizitätskonstante ϵ_r und einer Ladungsdichte ρ eingeführt. Die Ergebnisse für feste Werte der Spitzenparameter und festen Abstand illustriert Abbildung 3.9.

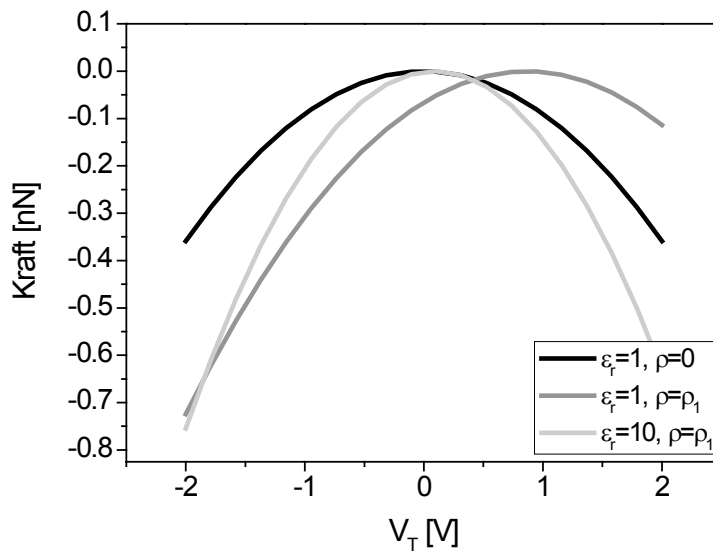


Abbildung 3.9: Simulierte Kraft-Spannungs-Kurven für das um ein Dielektrikum erweitertes System. Sowohl Kapazität der Schicht (variiert über die Dielektrizitätskonstante ϵ) als auch deren Ladungsdichte ρ haben deutlichen Einfluss auf messbare Größen.

Offenbar gibt es zwei wichtige Auswirkungen der Änderung dieser Werte: Zum Einen verschiebt sich die Position des Extremums der Kurven und die Krümmung der Kurven ändert sich. Verknüpft werden können genau diese Größen mit zwei der tatsächlichen Messung zugänglichen Größen, der von SKM ermittelten Spannung und dem von nc-SCM ermittelten Wert. Es scheint somit sinnvoll, diese Abhängigkeiten parametrisch zu untersuchen.

Da es sowohl aus numerischen als auch praktischen Gründen wenig zweckmäßig ist, die Dicke der Schicht als Charakteristikum zu nutzen, werden zwei Hilfsgrößen definiert:

$$\begin{aligned}\sigma &= \rho d \\ c &= \frac{\epsilon_0 \epsilon_r}{d}\end{aligned}\tag{3.2.1}$$

Die in (3.2.1) definierten Größen, die effektive Flächenladungsdichte σ und die Flächenkapazität c , beschreiben die dielektrischen Eigenschaften der Schichten vollständig. In Abbildung 3.10 sind die typischen Verläufe der Antwortfunktionen gezeigt. Die Variation der extrahierten Spannungen, welche sich im Bereich weniger Prozent bewegt, ist auf die Überlagerung kapazitiver und Coulombscher Wechselwirkung im System zurückzuführen. Die Änderung des extrahierten SCM-Signals ist hingegen verständlich, da gemäß (3.1.8) die Ableitung der Gesamtkapazität nach der Separation diesen Wert bestimmt.

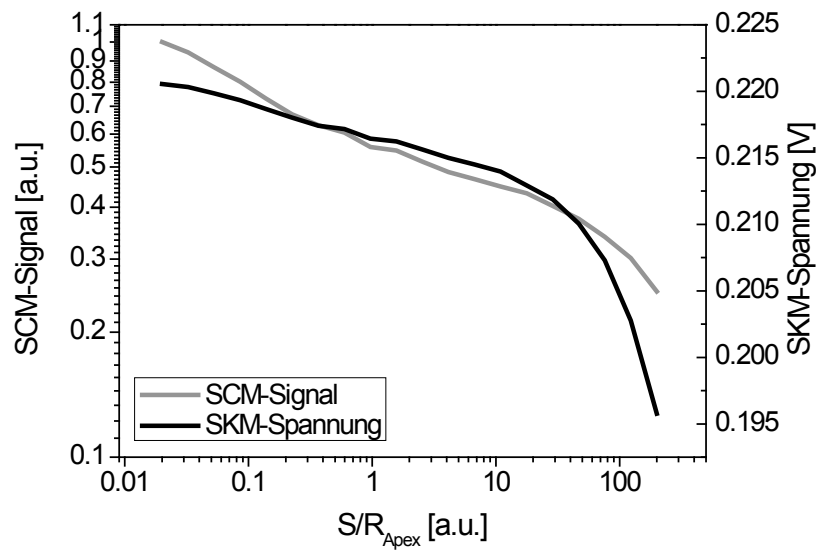


Abbildung 3.10: Simulierte Signal-Abstandsabhängigkeiten für SKM und SCM. Die in den SKM-Werten sichtbare Abhängigkeit erklärt sich durch die schwache Separationsabhängigkeit in (3.2.2). Der Verlauf des SCM-Signals ähnelt dem der Kraft-Abstands-Kurven ohne Dielektrikum.

Aus quasi eindimensionaler Betrachtungsweise der zugrunde liegenden Gleichungen ergeben sich für die erwartete SKM-Spannung U_{SKM} und das SCM-Signal A_{SCM} folgende Ausdrücke:

$$U_{SKM} = \frac{1}{2} \frac{\sigma}{c} - \frac{\sigma}{\epsilon_0} S_{eff} = \frac{1}{2} \frac{\sigma}{c} + \frac{\sigma}{\epsilon_0} \frac{C_{TS}}{\frac{\partial C_{TS}}{\partial S}} \quad (3.2.2)$$

$$A_{SCM} \sim \frac{c^2 \frac{\partial C_{TS}}{\partial S}}{(c + c_{TS})^2}$$

So ist zu erwarten, dass die gemessene Spannung maßgeblich von den Schichtparametern abhängt, während in das SCM-Signal A_{SCM} sowohl die Kapazität Spitze-Oberfläche der Schicht (c_{TS}) als auch deren Ableitung nach der Separation mit einfließt. Obwohl die formelmäßige Modellierung der gesamten Abhängigkeit, zumindest näherungsweise, möglich ist, sollen im Weiteren lediglich Sensitivitätskurven für übliche Separationen besprochen werden. Dazu wird zunächst bei konstanter Ladungsdichte die Dielektrizitätskonstante für verschiedene Separationen verändert. Die Ergebnisse dieser Prozedur zeigt Abbildung 3.11.

Offenbar lassen sich die extrahierten Werte gut an die Vorhersagen anpassen, eine günstige Vorschrift für ein Experiment lässt sich allerdings noch nicht herleiten. Während nämlich die Spannungen tatsächlich einer einfachen Formel folgen, was im Resultat auf einfache Messmöglichkeit schließen lässt, ist das für das Kapazitätssignal nicht der Fall. Allerdings kann man, wie in Experimenten üblich, versuchen, relative Signaländerungen über relativen Änderungen der interessanten Größen, hier der Dielektrizitätskonstante, darzustellen. Diese Möglichkeit wird in Abbildung 3.12 dargestellt.

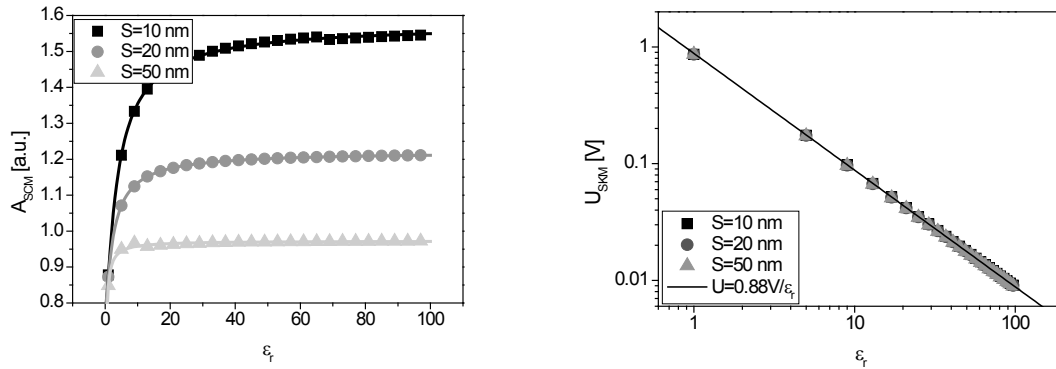


Abbildung 3.11: Sensitivitätskurven bezüglich Kapazität (links) und Ladung (rechts) in Abhängigkeit der relativen Dielektrizitätskonstante für verschiedene Separationen ($R_{Apex}=50$ nm, $d=100$ nm). Wie deutlich erkennbar ist, lassen sich die modellierten Werte (Punkte) gut an die Vorhersagen (Linien) anpassen. Des Weiteren ist ersichtlich, dass im dargestellten Separationsbereich, welcher im Übrigen den experimentell genutzten repräsentiert, das SKM-Signal nicht von der Separation abhängt.

Augenscheinlich verhält sich die relative Signaländerung proportional zum Quadrat der relativen Änderung der Dielektrizitätskonstante, wenn die Berechnung punktweise durchgeführt wird (linke Darstellung). Ein ähnliches Bild erhält man, wenn man die relative Änderung zu einem festen Bezugswert benutzt, wie es in einer Messung der Fall wäre (rechte Darstellung). Daraus folgt, dass tatsächlich im Rahmen eines Experimentes durch Kraftmessung Änderungen der dielektrischen Eigenschaften abgebildet werden können, auch mit zufriedenstellender Sensitivität. Allerdings stellt die schwer greifbare Spitze-Probe-Kapazität ein Problem dar, weshalb absolute Aussagen schwierig sind.

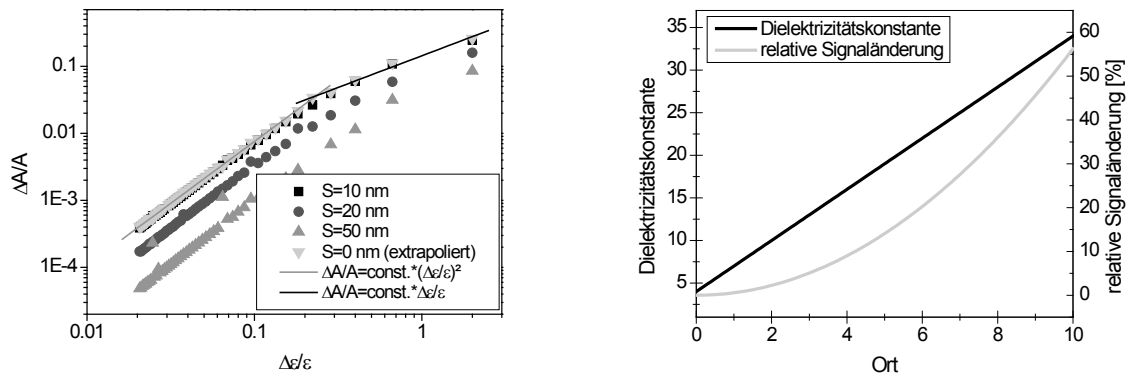


Abbildung 3.12: (Links) Relative Änderung des kapazitiven Signals über relativer Änderung der Dielektrizitätskonstante. Alle extrahierten Werte (Symbole) folgen einem typischen Verlauf, der bei Extrapolation der Separation auf null am deutlichsten wird. Für kleine Änderungen im Bereich von etwa 10% ist die zugehörige relative Signaländerung proportional zum Quadrat, für größere linear. (Rechts) Formelmäßig simulierter Scan über einen Bereich veränderlicher Dielektrizitätskonstante.

Ändert man bei konstanter Dielektrizitätskonstante (und Dicke) die Ladungsdichte der Schicht, ergeben sich die in Abbildung 3.13 illustrierten Verläufe. Eine Inspektion der dargestellten Verläufe der SCM-Signale liefert die Erkenntnis, dass im Rahmen der Modellierung die Ladungsdichte der Schichten keinen Einfluss auf kapazitive Messungen hat.

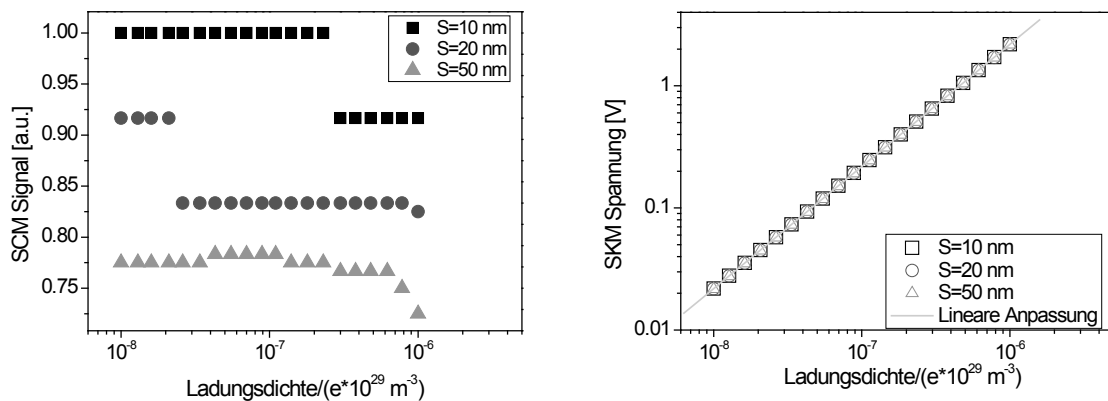


Abbildung 3.13: Verlauf des kapazitiven Signals (links) und der SKM Spannung in Abhängigkeit von der Ladungsdichte der Schicht für verschiedenen, sinnvollen, Separationen S . Bis auf numerische Fehler bleibt das Kapazitätssignal jeweils konstant, während die extrahierten Spannungen für alle Separationen direkt proportional zur Ladungsdichte sind.

Die präsentierten Ergebnisse legen zudem nahe, dass der separationsabhängige Term in (3.2.2) unter üblichen Bedingungen vernachlässigt werden kann, da die Variation der Kurven der SKM-Spannungen (rechts) für verschiedene Höhen gering ist. Offenbar wird somit, unabhängig von der Separation, immer das Verhältnis von Flächenladungsdichte und Flächenkapazität der dielektrischen Schicht bestimmt.

3.3 Untersuchungen zur Auflösung in einfachen Systemen

Neben Definition und Abschätzung der systemischen Transferfunktion (Kapitel 4) soll die räumliche Auflösung im Kontext dieser Arbeit anhand mehrerer Kriterien untersucht und festgelegt werden. Diese sind:

1. Minimale Größen in Zylindersymmetrie

Damit sind diejenigen Radien einer radialsymmetrischen Potentialverteilung gemeint, die eine definierte Änderung des durch SKM gemessenen Potentials erzeugen.

2. Kantenabbildung

Dazu wird die Spitze, oder Teile davon, in drei Dimensionen modelliert und über eine Potentialstufe geführt. Eine Inspektion des abgebildeten Verlaufs kann für Aussagen über die Auflösung genutzt werden.

3. Spektrale Dämpfung

Es wird untersucht, wie sinusförmige Potentialverteilungen abgebildet werden. Aufgrund der Möglichkeit der spektralen Zerlegung von zu messenden Signalen können daraus direkte Rückschlüsse auf die Auflösung gewonnen werden.

3.3.1 Radialsymmetrische Verteilung

Die intuitivste Möglichkeit, Informationen über das Auflösungsvermögen zu erhalten ist die Untersuchung der Abbildung einfacher Potentialverläufe. Beispielhaft hierzu ist eine radialsymmetrische Stufe, bei der das Potential der Ebene $z=0$ sprunghaft an einem bestimmten Wert $r=w$ von eins auf null abfällt.

Offenbar sollte das gemessene Potential ebenfalls zwischen null und eins variieren, abhängig von der Weite w der Stufe. Nimmt man zum Beispiel an, dass die Spitze des Systems über einen gewissen Radius R mittelt, ergibt sich durch Annahme einfacher Mittelwertbildung, also bei gleichwertigem Einfluss aller Oberflächenelemente:

$$U_{SKM} = \int_0^R \frac{2rV(r)}{R^2} dr = \begin{cases} \frac{w^2}{R^2}, & w < R \\ 1, & w \geq R \end{cases} \quad (3.3.1)$$

Das Ergebnis der bereits besprochenen numerischen Prozedur angewendet auf eine AFM-Spitze mit üblichen Parametern und einem Radius von $R_{Apex}=10$ nm zeigt Abbildung 3.14a. Zusätzlich zur Weite w der Potentialverteilung wurden verschiedenen Separationen S eingestellt. Offenbar folgt der minimal detektierbare Radius, willkürlich festgelegt durch gewisse Schwellspannungen, einfachen Potenzgesetzen, da die Konturlinien - die Linien gleicher gemessener Spannung - im dargestellten Plot nahezu gerade sind. Eine Linie der Form $w=S$ ist der Übersicht halber mit angegeben.

Die in Abbildung 3.14b extrahierten Schwellradien (Symbole) unterstützen diese Aussage, das Skalierungsverhalten erscheint quasi unabhängig vom gewählten Wert der Schwellspannung. Eine als universell angenommene Kurve (Linie) weicht dabei sehr schwach von einer linearen Funktion ab.

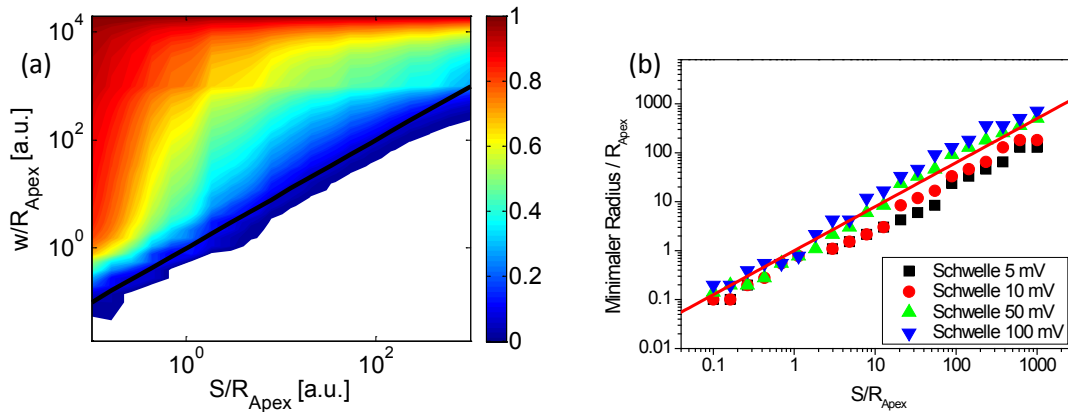


Abbildung 3.14: (a) Simulierter Verlauf der SKM-Spannung für radiale Stufen. Augenscheinlich folgt der minimal detektierbare Radius einfachen Potenzgesetzen, die ebenfalls angegebene Linie $w=S$ dient der Veranschaulichung. Extrahiert man aus der Matrix der Daten diese Werte w , an denen das gemessene Potential eine Schwelle überschreitet, ergibt sich (b).

Interessanterweise macht es bei dieser Auswertung keinen signifikanten Unterschied, wenn anstatt der Kraftwirkung auf die gesamte Spitze lediglich die auf den Apex wirkende Komponente analysiert wird, wie beispielhaft in Abbildung 3.15 (links) dargestellt ist. Einleuchtend erscheint dieses Verhalten bei Betrachtung der Verteilung der Energiedichte um eine AFM-Spitze (Abbildung 3.15 rechts): Die Energie des elektrischen Feldes, und damit die Wechselwirkung, ist am Apex konzentriert.

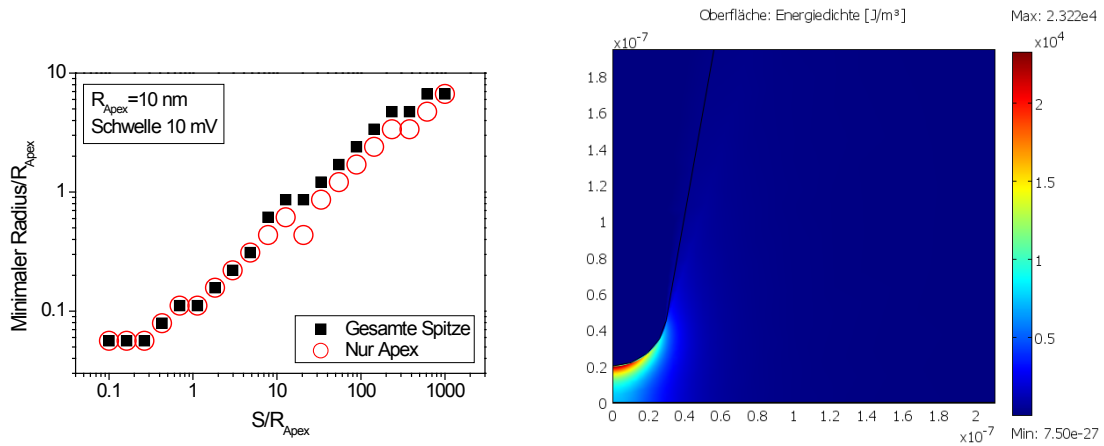


Abbildung 3.15: (Links) Bestimmung des minimalen Radius für eine Schwelle von 10 mV bei Betrachtung der gesamten Spitze (Quadrate) und des Apex (Kreise). Offenbar unterscheiden sich die Werte für kleine relative Separationen kaum, was darauf schließen lässt, dass der Apex der limitierende Faktor ist. Ebenfalls dargestellt (rechts) ist die Verteilung der Energiedichte im System, welche tatsächlich um den Apex konzentriert erscheint.

Die gezeigten Ergebnisse illustrieren dabei sehr anschaulich, dass der Apex der AFM-Spitze offenbar der limitierende Faktor in Messungen ist, da die extrahierten Daten für das Teilmodell nur wenig von denen des gesamten Modells abweichen. Offenbar kann man also annähernd davon ausgehen, dass sich das System unter üblichen Bedingungen gut über die elektrostatische Wechselwirkung mit einer Kugel beschreiben lässt.

Die Beschreibung der Auflösung wird häufig über die Angabe einer Transferfunktion durchgeführt [20], nähere Ausführungen dazu finden sich in Kapitel 4. An dieser Stelle sollen lediglich einige Konzepte genutzt werden, um die Ergebnisse zu modellieren. Die von SKM gemessene Spannung kann demnach mit der Potentialverteilung der Oberfläche über eine Faltung verknüpft werden:

$$U_{SKM} = \iint_A T(x, y) \cdot V(x, y) dA$$

$$T(x, y) = \frac{\frac{\partial c(x, y)}{\partial z}}{\int_A \frac{\partial c(x, y)}{\partial z} dA} \quad (3.3.2)$$

Die in (3.3.2) eingeführte Transferfunktion T enthält Informationen über die punktweise definierte Kapazität zwischen der Spitze und einem Ort auf der Oberfläche. Beachtet man die erwähnte Näherung des Systems als Kugel/Ebene und geht weiterhin davon aus, dass die besprochenen Kapazitäten nur vom Abstand d der Punkte zur Kugel abhängen, kann man angeben:

$$c \sim \frac{1}{d^p} \Rightarrow T(x, y, S) := \frac{p S^p}{2\pi (S^2 + x^2 + y^2)^{\frac{p}{2}+1}} \quad (3.3.3)$$

Die Annahme der Abhängigkeit zum Abstand potenziert mit einer Zahl p ist dabei durch die Multipolentwicklung des Kugel-Fläche-Problems (Anhang D) gegeben, in dem auch höhere

Potenzen auftreten. Mit Hilfe dieser Formel kann nun versucht werden, die simulierten Ergebnisse zu modellieren. Dazu muss lediglich das in (3.3.2) angegebene Integral berechnet werden:

$$U_{SKM} = 1 - \frac{S^p}{(w^2 + S^2)^{\frac{p}{2}}} \quad (3.3.4)$$

Die radiusabhängigen Werte aus Abbildung 3.14 sind in Abbildung 3.16 zusammen mit Modellkurven gemäß (3.3.4) dargestellt. Die Kurvenanpassungen zur Bestimmung der freien Parameter S und p liefern gute Übereinstimmungen, wenngleich die Werte für die Separation nicht der tatsächlichen entsprechen. Im speziellen Fall ergeben sich Exponenten p von etwa 2, statt der in der Simulation genutzten Separationen scheinen Werte in der Nähe der geometrischen Mittel von S und dem Radius des Apex R_{Apex} zu wirken.

Allerdings genügt das Wissen um ein brauchbares Modell nicht aus, um experimentelle Daten zu entfalten. Am günstigsten wäre dabei eine einfache Vorschrift der Art: *Bestimme den Messwert parametrisch und extrapoliere!* Für ein SKM-System wäre das, sofern möglich, die Messung der Spannung für unterschiedliche aber bekannte Abstände zur Oberfläche⁹.

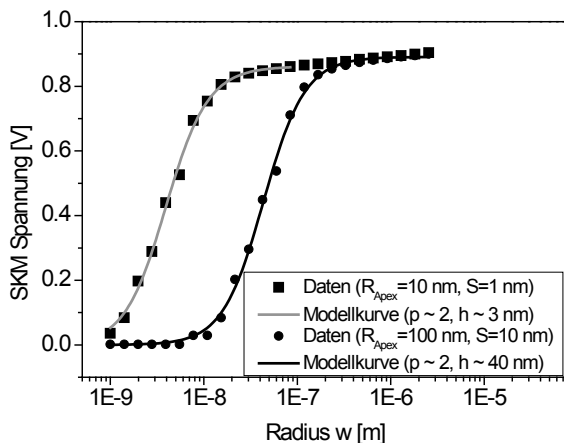


Abbildung 3.16: Modellierung der extrahierten Daten unter Annahme einfacher Abhängigkeiten. Die Übereinstimmung der Werte (Symbole) mit den Vorhersagen ist sehr gut. Allerdings ist zu beachten, dass nicht die tatsächliche Separation, sondern ein effektiver Wert genutzt werden muss.

Den Versuch einer einfachen Signalerückgewinnung illustriert Abbildung 3.17. Mit Hilfe simpler Potenzfunktionen gelingt die Entfaltung nicht, die extrapolierten Werte streuen innerhalb einer Größenordnung. Allerdings kann man unter bestimmten Bedingungen für die Strukturgrößen eine abgewandelte Formulierung benutzen, was zum gewünschten Ergebnis führt.

⁹ Im Falle des zweiphasigen Setups ist das ohne Weiteres möglich, bei simultaner Messung greift dieser Ansatz nicht (siehe Kapitel 2).

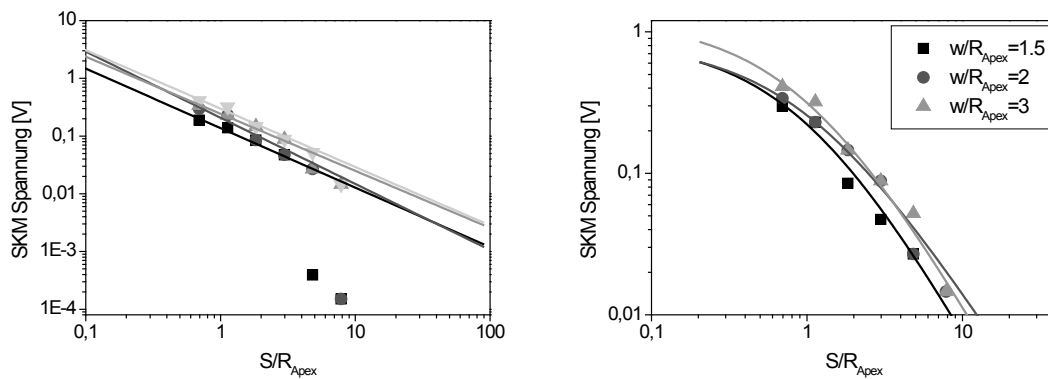


Abbildung 3.17: Versuch der Entfaltung der Messwerte für verschiedene Radien w mittels einfacher Potenzfunktionen (links). Die extrapolierten Werte streuen sehr stark, was eine sinnvolle Nutzung ausschließt. Allerdings ist es möglich, für Strukturen in der Größenordnung des Apex leicht abgewandelte Extrapolationsformeln zu nutzen, was bessere Ergebnisse liefert (rechts).

Die benutzte abgewandelte Form lautet:

$$U_{SKM} \sim \frac{1}{\left(1 + \frac{S}{R_{Apex}}\right)^2} \quad (3.3.5)$$

Eine Herleitung muss der Autor an dieser Stelle schuldig bleiben, motiviert ist das Vorgehen in einer Laurententwicklung von (3.3.4) um $S=R_{Apex}$ für Werte von $p=2$ und $w=R_{Apex}$.

3.3.2 Eindimensionale Verteilungen

Neben radialsymmetrischen Verteilungen, für die offenbar eine verlässliche Messung quantitativer Werte schwierig ist, sollen nun quasi eindimensionale Strukturen besprochen werden. Eindimensional bedeutet in diesem Kontext, dass eine der Ausdehnungen in x- oder y-Richtung deutlich grösser als die andere und zusätzlich grösser als der Einflussbereich der Messung ist. Einfache Untersuchungen wurden bereits veröffentlicht [21], die dargestellten Zugänge sind intuitiv und folgen den hier dargestellten Annahmen. Um im Gegensatz zu der im Artikel sehr vereinfacht angenommenen Geometrie allgemeinere Aussagen zu gewinnen, wird ein spektraler Ansatz gewählt: Es wird untersucht, wie die Amplitude $A=1$ einer sinusförmigen Potentialverteilung in Abhängigkeit der Separation gemessen werden kann. Dazu muss allerdings in kartesischer Geometrie gerechnet werden, was Komplexität und numerischen Aufwand deutlich erhöht.

Zur Bestimmung der Dämpfung wurden ebene Wellen der Form $V(x,y)=\cos(2\pi x/\lambda+\phi)$ unter der modellierten Spitze verschoben, dass heißt der Wert der Phase ϕ wurde geändert. Aus mehreren extrahierten Spannungen für eine Separation wurde dann mittels Kurvenanpassung an eine Sinusfunktion die Amplitude bestimmt. Dieses Vorgehen wurde für verschiedenen Separationen und Wellenlängen λ wiederholt. In Abbildung 3.18 sind Ergebnisse der Simulationen für verschiedene Radien des Apex (linkes Bild 10 nm, rechtes 50 nm) dargestellt. Das generelle Verhalten ähnelt sich, offenbar definiert R_{Apex} lediglich kritische Wellenlängen. Die stärkere Dämpfung kleinerer Wellenlängen entspricht dabei der allgemeinen Lebenserfahrungen eines

Wissenschaftlers (dieses Phänomen ist hierbei mit der Laplacegleichung verknüpft, es wird in Kapitel 4 besprochen).

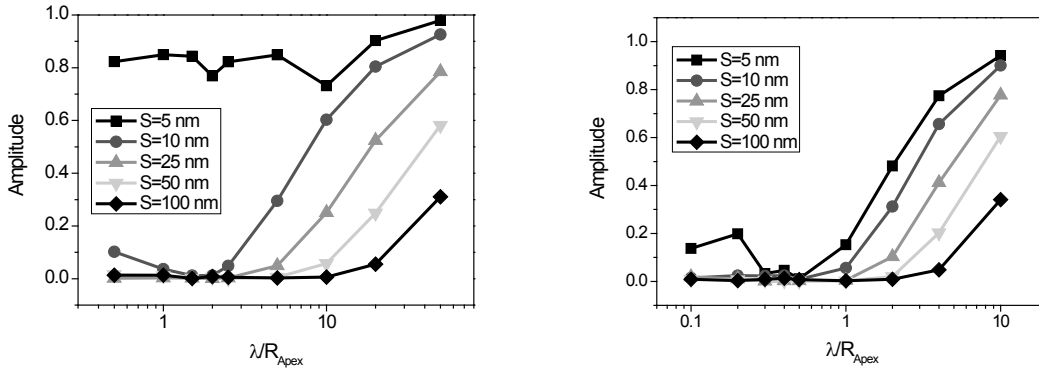


Abbildung 3.18: Spektrale Dämpfung bestimmt für unterschiedliche Separationen S und Radien des Apex (links $R_{Apex}=10$ nm, rechts $R_{Apex}=50$ nm). Für beide Radien deutlich erkennbar prinzipielle Verhalten der stärkeren Dämpfung kleiner Wellenlängen, wobei kritische Werte offenbar unterschiedlich sind.

Ausgehend von der zuvor definierten Transferfunktion kann nun eine Modellierung der Daten vorgenommen werden. Dazu muss beachtet werden, dass nun in kartesischen Koordinaten über eine Dimension integriert werden muss, um die eindimensionale Transferfunktion für die andere Koordinate zu erhalten. Integriert man also (3.3.3) über zum Beispiel „ y “, ergibt sich:

$$T(x) = \frac{\sqrt{\pi} p \Gamma\left(\frac{p+2}{2} - \frac{1}{2}\right) S^p (S^2 + x^2)^{\frac{1}{2} \frac{p+2}{2}}}{2\pi \Gamma\left(\frac{p+2}{2}\right)} \quad (3.3.6)$$

Diese etwas unhandliche, die Gammafunktion Γ enthaltende, Formulierung lässt sich für ganzzahlige Exponenten p etwas gefälliger darstellen:

$$\begin{aligned} p=1 &\Rightarrow T(x) = \frac{1}{\pi} \frac{S}{x^2 + S^2} \\ p=3 &\Rightarrow T(x) = \frac{2}{\pi} \frac{S^3}{(x^2 + S^2)^2} \end{aligned} \quad (3.3.7)$$

Die Transferfunktionen sind also unter den besprochenen Annahmen der Lorentzfunktion ähnliche Funktionen, von denen einige in Abbildung 3.19 dargestellt sind. Die Größe S definiert augenscheinlich die Breite der Graphen, womit sie gleichzeitig auch der entscheidende Parameter der Messung wird. Setzt man, wie gefordert, Sinusfunktionen in das Faltungsintegral ein, ergeben sich für $p=1$ und $p=3$ folgende Amplitudengänge:

$$\begin{aligned} p=1 &\Rightarrow A(S, \lambda) = e^{-\frac{2\pi S}{\lambda}} \\ p=3 &\Rightarrow A(S, \lambda) = 2\pi \frac{S}{\lambda} + e^{-\frac{2\pi S}{\lambda}} \end{aligned} \quad (3.3.8)$$

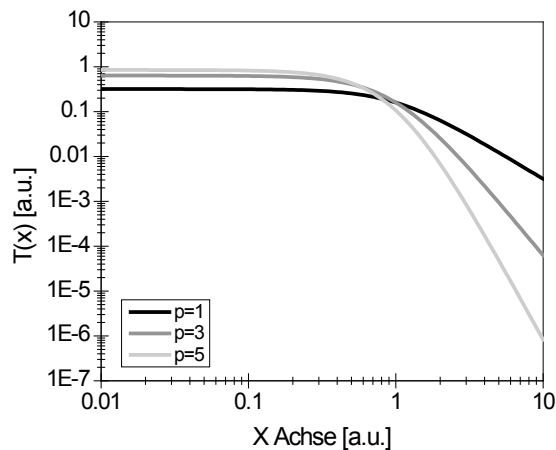


Abbildung 3.19: Graphen der eindimensionalen Transferfunktionen für verschiedenen Exponenten und feste Separation $S=1$. Die Zahl S bestimmt dabei, quasi unabhängig vom Exponenten p , die Halbwertsbreite der Funktion.

Motiviert von den in (3.3.8) gegebenen Gleichungen kann somit der Amplitudengang an eine Funktion folgender Form angepasst werden:

$$A = \left(1 + b \frac{cR_{Apex}}{\lambda} \right) e^{-\frac{cR_{Apex}}{\lambda}} \quad (3.3.9)$$

Die Ergebnisse dieser Anpassungen an Daten für $R_{Apex}=25\text{nm}$ und verschiedene Separationen S zeigt Abbildung 3.20 (links). Die Übereinstimmung der Modellkurven mit den Werten ist sehr gut, was für die Sinnhaftigkeit der Annahmen spricht. Der praktisch rein exponentielle Zusammenhang zwischen Amplitude und Separation erlaubt zudem eine einfache Entfaltung: Eine Darstellung des Logarithmus der gemessenen Amplituden über der Separation gestattet lineare Extrapolation auf $S=0$ und somit auf den tatsächlichen Wert $A=1$ (rechtes Bild).

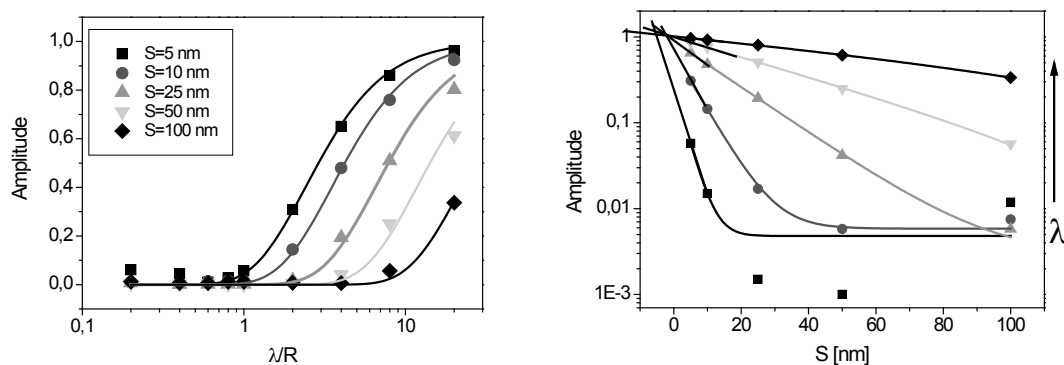


Abbildung 3.20: Anpassung der simulierten Werte an Modellfunktionen (links). Die gute Übereinstimmung der Daten mit der hauptsächlich exponentiellen Modellfunktion motiviert eine einfache Extrapolation der Werte auf $S=0$ (rechts). Die Entfaltung greift für die dargestellten Wellenlängen im Bereich einiger Vielfacher des Apexradius sehr gut.

Zusammenfassend kann formuliert werden:

1. Unter üblichen Messbedingungen limitieren der Abstand S zur Probe und der Radius des Apex die Messwerte im SKM-Modus.
2. Für radialsymmetrische Potentialverteilungen können die erwarteten Werte gut modelliert werden, eine einfache Signalentfaltung ist allerdings nicht möglich.
3. Für quasi eindimensionale Strukturen werden alle spektralen Komponenten exponentiell gedämpft, was eine einfache Entfaltung erlaubt. Zudem können gemessene Profile als Faltung zwischen tatsächlichen Profilen und Lorentzfunktionen aufgefasst werden und durch entsprechende Entfaltung die Amplituden der Oberflächenpotentialfunktionen bestimmt werden.

3.3.2 Auflösung kapazitiver Änderungen

Zusätzlich zu den Bemerkungen bezüglich der Auflösung in SKM-Setups sollen noch einige Aussagen zur örtlichen Auflösung bei Nutzung der kapazitiven Signale gemacht werden. Wie in Absatz 3.2 gezeigt, hängt die zweite Ableitung der Kraft, und damit das Signal im SCM-Modus, von der Kapazität einer dielektrischen Zwischenschicht ab. Interessant ist allerdings die Frage, wie groß die laterale Ausdehnung dieser Schicht sein muss, um einen Unterschied im Signal zu bemerken. Alternativ dazu ist es wünschenswert, für einfache Abschätzungen einen Mittelungsradius angeben zu können.

Zur numerischen Abschätzung dieser Größen wurde die Dielektrizitätskonstante moduliert, hier sprungförmig fallend an einem gewissen Radius r_0 . In Abhängigkeit dieses Radius ergibt sich eine Darstellung gemäß Abbildung 3.21. Es zeigt sich offensichtlich eine Änderung des SCM-Signals A_{SCM} , wenn sich der Sprungradius im Bereich des Apexradius bewegt, für Werte grösser als zehnmal der Apexradius bleibt der Wert stabil. Vorsichtig kann also formuliert werden, dass der Mittelungsradius im Bereich kleiner Vielfache des Apexradius liegt, also eine mittlere Kapazität aus diesem Bereich bestimmt wird. Geht man zur Modellierung der Ergebnisse zum Beispiel von (3.3.2) aus, kann man ansetzen:

$$\frac{d^2 F}{dU^2} \sim A_{SCM} \sim \int_0^\infty r \cdot c(r) \frac{\frac{\epsilon_0 S}{(S^2 + r^2)^{\frac{3}{2}}}}{\left(c(r) + \frac{\epsilon_0}{\sqrt{S^2 + r^2}} \right)^2} dr \quad (3.3.10)$$

Ausführen des Integrals und Normierung der Werte für Separation S , Dicke D der Schicht und des Radius r_0 auf R_{Apex} führt zu:

$$\frac{A_{SCM}(a)}{A_{SCM}(\infty)} = \frac{1}{1 + \frac{d}{\sqrt{s^2 + a^2}}} - \frac{1}{1 + \frac{d}{s}} + f_0; a = \frac{r_0}{R_{Apex}}, d = \frac{D}{\epsilon_r R_{Apex}} \quad (3.3.11)$$

Modellkurven gemäß dieser Vorschrift sind ebenfalls eingefügt, die Übereinstimmung ist überaus gut. Man kann demnach die entwickelte Formel nutzen, um Faustregeln zu entwickeln. Sucht man so zum Beispiel in (3.3.11) den Punkt, an dem die Signaländerung die Hälfte des maximalen Wertes erreicht, ergibt sich $a \approx 2s$. Wie den vorigen Überlegungen zu entnehmen ist,

ergab sich die Auflösung eines SKM-Experiments ebenfalls als bestimmt zur relativen Separation (3.3.4), ebenso wie die eingeführte effektive Fläche der AFM-Spitze (3.1.13).

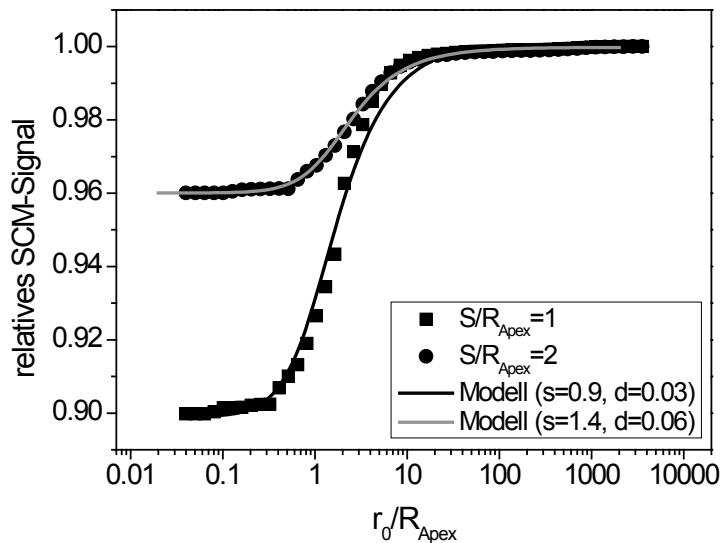


Abbildung 3.21: Entwicklung des relativen SCM-Signals für eine radiale Stufe in der Dielektrizitätskonstante für verschiedene Separationen. Die numerisch erhaltenen Daten (Symbole) suggerieren Auflösung im Bereich des Apexradius. Das einfache Modell (Linien) erlaubt gute Nachstellung, wobei die effektiven Separationen s und Dicken d tatsächlich im erwarteten Bereich liegen.

Somit kann geschlussfolgert werden, dass für ein SCM-Experiment die Separation zwischen Spitze und Probe zwei Einflüsse hat: Zum einen ändern sich die gemessenen Signale aufgrund der Serienschaltung von Proben- und Spitzenkapazität und zum anderen steigt der Mittelungsradius mit der Separation. Dieser Mittelungsbereich hat eine Ausdehnung in der Größenordnung der Separation.

Weitere nützliche Abschätzungen erhält man für den Fall quasi eindimensionaler Verteilungen von Oberflächenkapazitäten. Beachtet man, dass übliche Oxidkapazitäten und Kapazitäten von Inversionsschichten in der Größenordnung einiger Zehnfacher (Faktor f) der Spitze-Probe-Kapazität liegen, kann man ansetzen:

$$c_{\text{Oberfläche}} = f c_{\text{Spitze-Oberfläche}} + \Delta(x)$$

$$c_{\text{Oberfläche}}^2 \frac{\frac{\partial}{\partial z} c_{\text{Spitze-Oberfläche}}}{(c_{\text{Oberfläche}} + c_{\text{Spitze-Oberfläche}})^2} \approx \frac{\partial}{\partial z} c_{\text{Spitze-Oberfläche}} + 2 \frac{\frac{\partial}{\partial z} c_{\text{Spitze-Oberfläche}}}{f \cdot c_{\text{Spitze-Oberfläche}}} \Delta(x) \quad (3.3.12)$$

Dabei wurde die Variation der mittleren, punktweise betrachteten, Oberflächenkapazität als Δ bezeichnet. Einsetzen in die integralen Formulierungen ergibt:

$$A_{\text{SCM}} \sim \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial}{\partial z} c_{\text{Spitze-Oberfläche}} dx dy + \int_{-\infty}^{\infty} \Delta(x) \left(\int_{-\infty}^{\infty} 2 \frac{\frac{\partial}{\partial z} c_{\text{Spitze-Oberfläche}}}{f \cdot c_{\text{Spitze-Oberfläche}}} dy \right) dx \quad (3.3.13)$$

Unter Annahme bestimmter, punktwiser, Spitze-Oberfläche-Kapazitäten und einer harmonisch variierenden Funktion Δ erhält man:

$$A_{SCM} \sim A_{SCM,0}(z) + \int_{-\infty}^{\infty} \frac{A}{f} \cos(kx) \frac{\pi z}{\sqrt{z^2 + x^2}} dx = A_{SCM,0}(z) + \frac{A}{f} \pi z K_0(kz) \quad (3.3.14)$$

Die in (3.3.14) eingeführte Funktion K_0 ist die sogenannte modifizierte Besselfunktion zweiter Art der Ordnung null. Sinnvolle Approximationen ergeben sich für große kz (also Separationen z grösser als etwa die halbe Wellenlänge) und kleine kz zu:

$$A_{SCM} \sim A_{SCM,0}(z) + \begin{cases} \frac{A}{f} \pi^{\frac{3}{2}} \sqrt{z} e^{-kz} & \text{große } kz \\ \frac{A}{f} \pi z \left(\ln\left(\frac{2}{kz}\right) - 0.57 \right) \frac{1}{\sqrt{z}} & \text{kleine } kz \end{cases} \quad (3.3.15)$$

Da die (effektiven) Separationen im Bereich des Apexradius liegen, sollte für sinnvolle Messung (und auch Gültigkeit der angesetzten Näherung) immer der Fall für große kz gelten, prinzipiell wird also eine Variation in der Oberflächenkapazität exponentiell mit der Höhe gedämpft.

Zur numerischen Überprüfung wurde die Dielektrizitätskonstante einer Schicht sinusförmig variiert und der Wert der Krümmung der Kraft-Spannungs-Kurve in Abhängigkeit von Separation S und Wellenlänge λ bestimmt. Die Rechnungen wurden aufgrund der Komplexität auf den Apex limitiert und für wenige Wellenlängen und Separationen durchgeführt. Zur Extraktion der Parameter wurde die Spitze virtuell über die Probe gescannt und Mittelwert beziehungsweise Amplitude der gemessenen Schwankung bestimmt. Eine Darstellung dieser virtuellen Scans für feste Separation ($S/R_{Apex}=0.1$) und verschiedene Wellenlängen zeigt Abbildung 3.22a. Offensichtlich lassen sich die Messwerte an Cosinusreihen der Form $A \cos(x/\lambda) + B \cos^2(x/\lambda)$ anpassen, eine Darstellung der Amplituden über den Wellenzahlen zeigt Abbildung 3.22b.

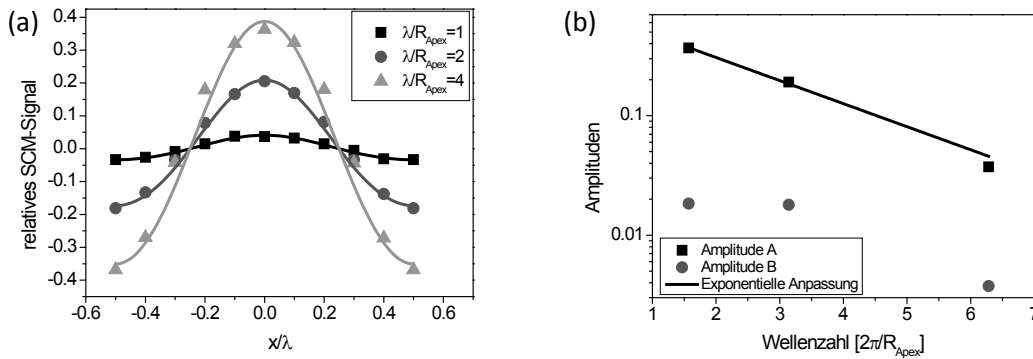
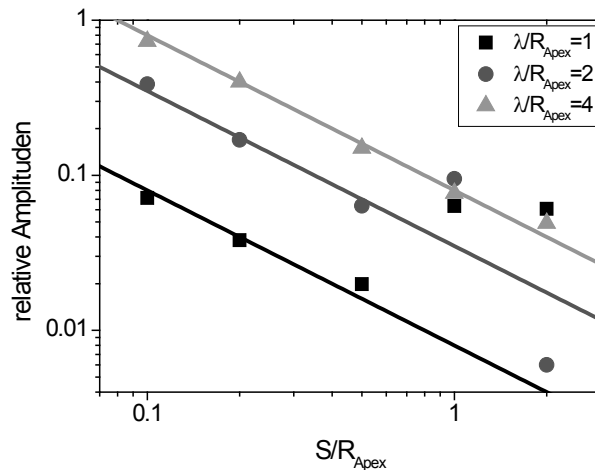


Abbildung 3.22: Simulierte SCM-Scans entlang kapazitiver Inhomogenitäten. Die extrahierten Werte (Symbole), hier die relative Amplitude, wurden an eine Cosinusreihe angepasst (Linien). Die relativen Amplituden variieren sich dabei exponentiell mit der Wellenzahl (rechts).

Für feste Separationen lassen sich die Amplituden demnach an eine exponentielle Funktion der Wellenzahl anpassen, was der Abschätzung in (3.3.15) entspricht. Trägt man allerdings die Separationsabhängigkeit der Amplituden auf, ergibt sich im Gegensatz zu den Erwartungen eine Abhängigkeit in Form eines Potenzgesetzes. Daraus kann gefolgert werden, dass sowohl eine Mittelung, welche exponentielle Dämpfung erzeugt, als auch die prinzipiell bedingte Schwächung

der Signale eine Rolle spielen. Es darf also davon ausgegangen werden, dass auch SCM Auflösung im Bereich von SKM erbringt, die Signale allerdings eher durch die starke Separationsabhängigkeit der Gesamtkapazität bestimmt sind, was sich in Gleichung (3.3.15) durch Anstieg des Verhältnisses f einbringen lässt.



$$A(s, k) = \frac{\exp(-ks)}{(1+s)^p} \quad (3.3.16)$$

Abbildung 3.23: Entwicklung der Amplituden mit der Separation. Allgemein gilt, dass kleinere Wellenlängen stärker gedämpft werden und die (relative) Amplitude einen potenziellen Zusammenhang mit der Separation aufweist. Dieser Fakt weist darauf hin, dass laterale Mittelung der prinzipiellen Dämpfung untergeordnet ist. Somit sollte die Abhängigkeit der Amplitude von den Wellenzahlen und Separationen die Form (3.3.16) annehmen.

3.4 Komplexe Systeme – Metallische Spitze / Halbleiteroberfläche

Im vorhergehenden Abschnitt wurde gezeigt, dass die Kraft zwischen einer metallischen Sonde und einer metallischen Oberfläche relativ einfach mittels kapazitiver und Coulombwechselwirkung beschrieben werden kann. Im Fall einer halbleitenden Probe ist die Situation allerdings komplizierter, da die Ladungsträgerdichte nahe der Oberfläche stark vom elektrostatischen Potential und damit von der Spannung der Sonde abhängt.

In den Anhängen A1 und A3 werden grundlegende Gedanken der Elektrostatik in Halbleitern und deren Verhalten unter Beleuchtung dargestellt, auf eine detaillierte Beschreibung wird daher hier verzichtet.

Zur numerischen Simulation der Wechselwirkung zwischen Spitze und Probe wird die Domäne um ein Gebiet erweitert, in dem die Poissongleichung für Silizium (Gleichung A1.4) gelöst wird (Abbildung 3.24). Die Randbedingung am Übergang zwischen dem Halbleiter und Luft wird zunächst als kontinuierlich angenommen. Um die Differenz der Austrittsarbeiten von metallischer Spitze und Halbleiter zu modellieren, wird das Potential an der Spitzensoberfläche um die Flachbandspannung verringert. Die Berechnungen werden ebenfalls in Zylindersymmetrie ausgeführt, um die Anzahl der Freiheitsgrade zu verringern.

Durch Integration des Maxwellschen Stresstensors über Rand „1“ ergibt sich wiederum die auf die Spitze wirkende Kraft. Da eine Abhängigkeit der Kraft vom Wert der angelegten Spannung (V_T) zu erwarten ist, wird die Berechnung parametrisch durchgeführt. Ein Beispielergebnis zeigt Abbildung 3.25. Sehr deutlich zeigt sich in dieser Darstellung der Einfluss der Dotierung: Zum einen verschiebt sich die Lage des Kraftminimums und zum anderen ändert sich die

Krümmung der Kurven. Diese beiden Werte sind, wie bereits beschrieben, mit der Messung zugänglichen Parametern verknüpft, nämlich der SKM-Spannung und dem SCM-Signal.

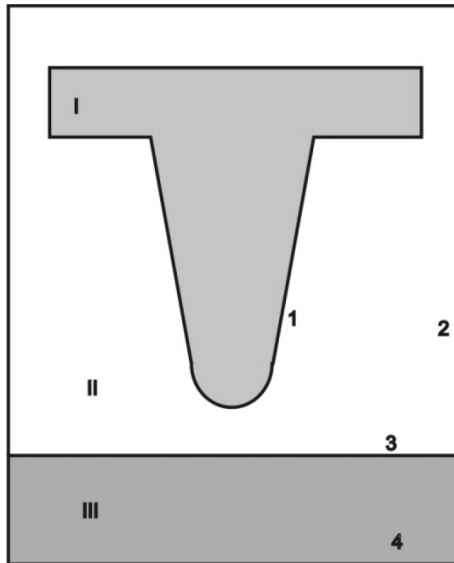


Abbildung 3.24: Erweitertes Gebiet für die Modellierung der Kraftwirkung zwischen metallischer Spitze und Halbleiter. Im zusätzlichen Gebiet III muss die nichtlineare Poissongleichung für Silizium gelöst werden.

Die Form der Kurven in Abbildung 3.25 ist im Großen und Ganzen noch als parabolisch zu bezeichnen, obwohl, verglichen mit der Kraft-Spannungs-Kurve für ein Spitze-Metall System, die Graphen asymmetrisch sind und sich sowohl lokale Steigung und Krümmung ändern (Abbildung 3.26).

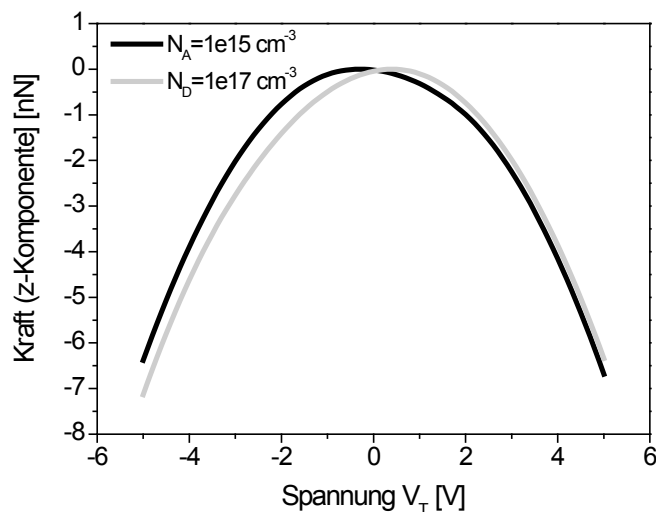


Abbildung 3.25: Typische Kraft-Spannungskurven für den Fall einer metallischen Spitze über unterschiedlich dotiertem Material. Offenbar hat der Wert der Dotierung Einfluss sowohl auf die Position des Extremums als auch auf die Krümmung der Kurven.

Grundlegende Wechselwirkungen im System Spitze-Probe

Um den Verlauf der Kurven zu erklären, soll zunächst eine eindimensionale Betrachtung benutzt werden. Die im System auf die Spitze wirkende Kraft berechnet sich im Rahmen dieser Näherung wie folgt:

$$F_{Tip} = \frac{A}{2} \cdot \epsilon_0 \cdot E_{Tip}^2 \quad (3.4.1)$$

Dabei ist A die Fläche, hier nur als einheitenbehafteter Proportionalitätsfaktor zu betrachten, und E_{Tip} das an der Spitze wirkende elektrische Feld.

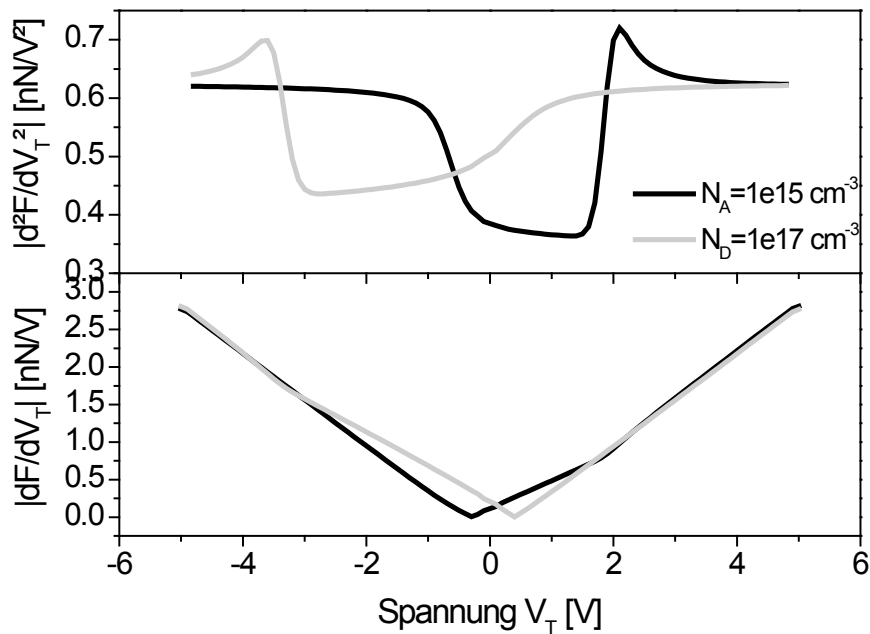


Abbildung 3.26: Aus den Daten in Abbildung 3.25 berechnete Verläufe von Steigung und Krümmung. Die bereits erwähnten Unterschiede treten in diesen Darstellungen noch deutlicher hervor.

Bildet man die Ableitung von (3.4.1) nach der Spannung, erhält man:

$$\begin{aligned} \frac{\partial F_{Tip}}{\partial V_T} &= A \cdot \epsilon_0 \cdot E_{Tip} \cdot \frac{\partial E_{Tip}}{\partial V_T} \\ \frac{\partial^2 F_{Tip}}{\partial V_T^2} &= A \cdot \epsilon_0 \cdot \left(E_{Tip} \cdot \frac{\partial^2 E_{Tip}}{\partial V_T^2} + \left(\frac{\partial E_{Tip}}{\partial V_T} \right)^2 \right) \end{aligned} \quad (3.4.2)$$

Wenn man beachtet, dass das lokale elektrische Feld an der Oberfläche eines Metalls mit der Oberflächenladungsdichte verknüpft ist, kann Gleichung (3.4.2) umgewandelt werden in:

$$\begin{aligned} \frac{\partial F_{Tip}}{\partial V_T} &= Q_{Tip} \cdot \frac{\partial Q_{Tip}}{\partial V_T} \cdot \frac{1}{\epsilon_0 \cdot A} \\ \frac{\partial^2 F_{Tip}}{\partial V_T^2} &= \frac{1}{\epsilon_0 \cdot A} \cdot \left(Q_{Tip} \cdot \frac{\partial^2 Q_{Tip}}{\partial V_T^2} + \left(\frac{\partial Q_{Tip}}{\partial V_T} \right)^2 \right) \end{aligned} \quad (3.4.3)$$

Vergleicht man nun obige Gleichung mit den bekannten Größen aus MOS-Systemen, erkennt man schließlich, dass die Steigung der Kraft-Spannungs-Kurve proportional zum Produkt aus Ladung auf der Spitze und differentieller Kapazität ist. Da die Ladung in Akkumulation und Inversion proportional zur angelegten Spannung ist (die Halbleiteroberfläche verhält sich wie ein Metall, siehe zum Beispiel [22]) und die differentielle Kapazität in diesen Bereichen konstant ist (siehe Abbildung A1.3 in A1), ergibt sich ein linearer Verlauf der Steigungs-Spannungs-Kurve. Im Bereich der Verarmung, hier zwischen etwa -1 V und 1 V, erhält man aufgrund der geringeren Kapazität eine verringerte Steigung. Der Nullpunkt der Steigungskurve ergibt sich, wenn sich keine Ladung auf der Spitze befindet. Dieser Punkt ist erreicht, wenn die Spannung gleich der Flachbandspannung V_{FB} ist, weil dann das elektrische Feld an der Halbleiteroberfläche und damit im gesamten Raum zwischen Oberfläche und Spitze verschwindet. Da die Flachbandspannung ohne Einwirkung externer Ladungen die Differenz der Austrittsarbeiten von Spitze und Halbleitermaterial ist, ist der Wert der Spannung für beide dargestellten Kurven unterschiedlich.

Deutlichere Abweichungen zwischen den Dotierungen ergeben sich in der Krümmungs-Spannungs-Kurve: Wie von Gleichung (3.4.3) nahe gelegt, ähneln die Graphen stark Kapazitäts-Spannungs-Kurven von MOS-Systemen und spiegeln deren prinzipiellen Eigenschaften. So ist es zum Beispiel möglich, die Art der Dotierung zu erkennen, da die Kapazität, und damit auch die Krümmungskurve, üblicherweise in Akkumulation vom Betrag höher ist als in Verarmung. Da das Halbleitermaterial bei angelegter Spannung vom jeweils gleichem Vorzeichen „verarmt“ (also positive Spannung bei p-leitendem Material), kann die graue Kurve als zu n-leitendem Material gehörig identifiziert werden. Des Weiteren deuten die Unterschiede in sowohl minimalem Wert der Krümmung als auch in der Spannung, bei der der Wert der Krümmung wieder zunimmt, darauf hin, dass auch die Konzentrationen der Dotierstoffe unterschiedlich sind. Dieses Verhalten ist in der Tatsache begründet, dass die maximale Weite der Verarmungszone stark mit der Dotierung abnimmt (vgl. A1) und so die minimale Kapazität zunimmt.

Eine genauere Untersuchung der besprochenen Abhängigkeiten ergibt folgende Resultate: Mit Erhöhung der Dotierkonzentration nimmt die Asymmetrie der Steigungs-Spannungs-Kurven (Abbildung 3.27a) ab und die Position des Minimums verschiebt sich (b). Die Verschiebung ist in der unterschiedlichen Austrittsarbeit des Siliziums begründet, was bei fester Austrittsarbeit des Spitzenmaterials zu einer Verschiebung der Flachbandspannung führt. Eine Darstellung der Minima über der Dotierungskonzentration zeigt den erwarteten, durch die Lage der Fermi-niveaus bestimmten, logarithmischen Verlauf. Die Krümmungs-Spannungs-Kurven werden mit zunehmender Dotierung flacher, das heißt, dass sich sowohl minimaler Wert als auch die Position des Wiederanstiegs ändern (Abbildung 3.28 links).

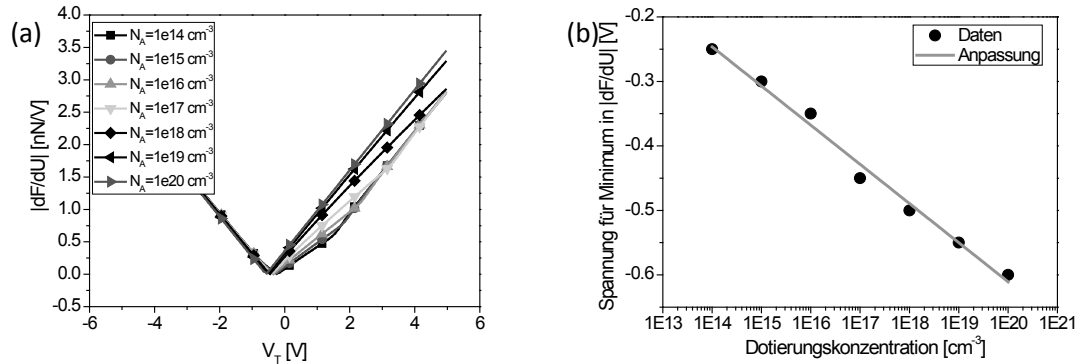


Abbildung 3.27: (a) Abhängigkeit der Steigungs-Spannungs-Kurve von der Dotierung für einen Spitze-Probe-Abstand von 5 nm. Sehr deutlich erkennbar ist eine Verschiebung der SKM-Spannung, welche wie vermutet logarithmisch mit der Dotierkonzentration skaliert (b).

Die Veränderungen der Krümmungskurve lassen sich mit Hilfe der Verarmungsnaheung erläutern: Wenn das Potential an der Oberfläche des Siliziums einen gewissen Wert erreicht, sind etwa genauso viele Minoritätsladungsträger vorhanden, wie Majoritätsladungsträger im Falle ohne äußeres Potential. Der zugehörige Wert der Spannung an der SPM-Spitze (bzw. am Gatekontakt eines MOS-Systems) wird Schwellspannung V_{Th} genannt. Die Abhängigkeit der Schwellspannung von der Dotierung N ist bei eindimensionaler Betrachtung gegeben durch:

$$V_{Th} = V_{FB} + 2 \cdot \phi_F + \frac{\sqrt{4 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot N \cdot \phi_F}}{C_{TS}} \quad (3.4.4)$$

In (3.4.4) stellt ϕ_F das sogenannte „Bulkpotential“ dar, welches vom Betrag her die Differenz der Lage des Fermi-niveaus und der Mitte des Bandgaps angibt. Eine Kurvenanpassung der Positionen des Wiederanstiegs unter Benutzung dieser Funktion, mit der effektiven Kapazität des Spitze-Oberfläche-Systems C_{TS} , ergibt eine gute Übereinstimmung der Werte aus der Simulation mit der Vorhersage (Abbildung 3.28 rechts). Die Beschreibung der Abhängigkeit der minimalen Krümmung kann mit Hilfe von Gleichung (3.4.3) erfolgen, wonach die zweite Ableitung der Kraft nach der Spannung proportional zum Quadrat der differentiellen Kapazität ist (der d^2Q/dV_T^2 enthaltende Term entfällt in Bereichen geringer Krümmung). Die minimale differentielle Kapazität C_{min} wiederum ergibt sich zu:

$$C_{min} = \frac{1}{\frac{1}{C_{TS}} + \frac{x}{A_{TS} \cdot \epsilon_0 \cdot \epsilon_{Si}}} \quad (3.4.5)$$

Darin ist A_{TS} eine effektive Fläche und x die maximale Weite der Verarmungszone, die gegeben ist durch:

$$x = \sqrt{\frac{4 \cdot \phi_F \cdot \epsilon_0 \cdot \epsilon_{Si}}{e \cdot N}} \quad (3.4.6)$$

Eine Anpassung der simulierten Werte an eine Funktion der Form C_{min}^2 ist leider nicht zufriedenstellend möglich. Wenn allerdings in Betracht gezogen wird, dass der Einfluss der Form

der Spitze eine zusätzliche, konstante, Krümmung verursacht, können die Werte gut reproduziert werden (Abbildung 3.28 rechts).

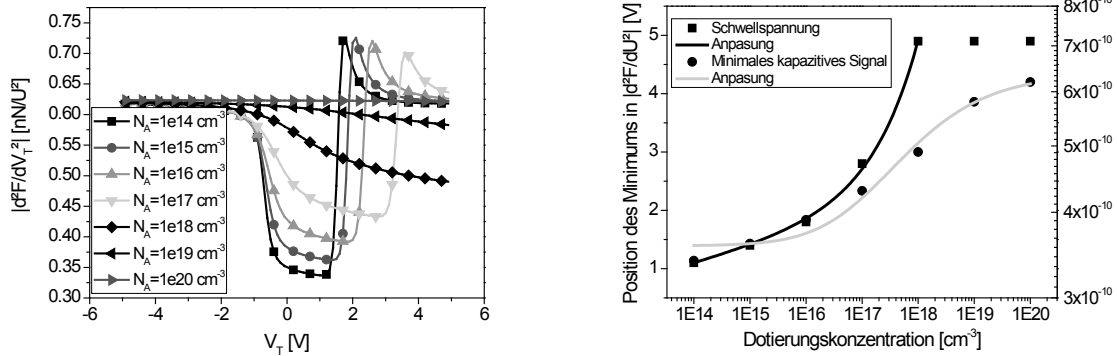


Abbildung 3.28: Abhängigkeit der Krümmung von der Dotierung für einen Spitze-Probe-Abstand von 5 nm (links). Sowohl Position des Minimums der Kurve („Schwellschwellspannung“) als auch der Wert an dieser Stelle lassen sich gut modellieren (rechts).

Eine Auswertung der Krümmung einer Kraft-Spannungs-Kurve erlaubt somit Aussagen über die lokale¹⁰ Dotierungskonzentration, wobei der zu erwartende minimale Wert näherungsweise proportional zur Konzentration ist. Unglücklicherweise ist der Einfluss der Spitze-Probe-Kapazität, maßgeblich bestimmt durch den Abstand zwischen Spitze und Probe, relativ groß, wie Abbildung 3.29 zeigt. Die Abhängigkeit des Signals ist dabei für alle Dotierungen etwa mit $A_{SCM} \sim S^{-0.6}$ gegeben.

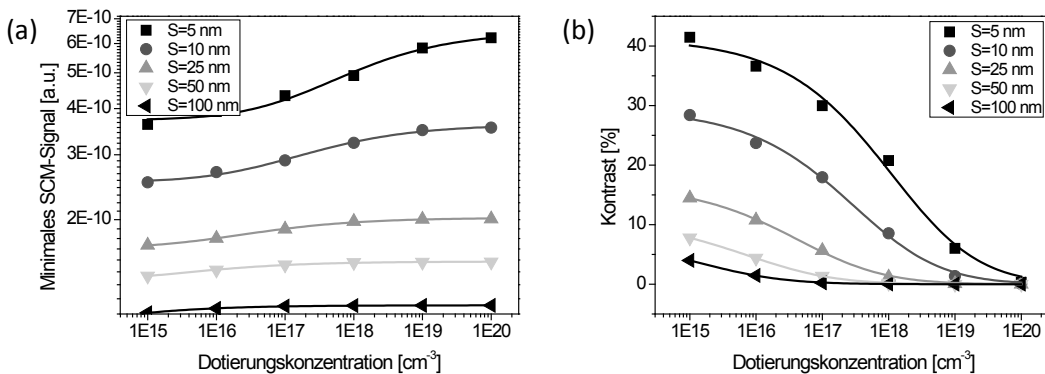


Abbildung 3.29: Abhängigkeit des SCM-Signals für verschiedene Separationen (Symbole) nebst Kurvenanpassungen (a) und (b) Kontrastfunktion gemäß (3.4.7).

Der in Abbildung 3.29b dargestellte Kontrast wurde wie folgt definiert: Vom maximalen Wert der Krümmung wurde der minimale Wert abgezogen und die Differenz durch den maximalen Wert geteilt. Durch Vergleich mit Gleichung (3.4.5) und Abbildung 3.28 ergibt sich somit folgende Kontrastfunktion KF :

¹⁰ Lokal bedeutet in diesem Zusammenhang zunächst mit Auflösung in der Größenordnung der Debyelänge des jeweiligen Materials. Wie zum Beispiel p-n-Übergänge das Signal beeinflussen, soll später geklärt werden.

$$KF = \frac{F_{\max} - F_{\min}}{F_{\max}} \approx \frac{x}{S} \cdot \frac{\frac{x}{S} + 2 \cdot \varepsilon_{r, Si}}{\left(\frac{x}{S} + \varepsilon_{r, Si}\right)^2} \quad (3.4.7)$$

Offensichtlich hängt der Kontrast lediglich vom Verhältnis der Weite der Verarmungszone zum Abstand zwischen Spitze und Probe ab. Das bedeutet, dass eine sinnvolle Abbildung nur für Abstände geringer als etwa das Zehnfache der Weite der Verarmungszone möglich ist. Gemäß Abbildung 3.30 muss also der Abstand im Bereich einiger 10 nm liegen, was allerdings lediglich einen geringen Kontrast für hoch dotierte Proben ergibt. Abbildung 3.30 kann auch zur Abschätzung der Signalsensitivität genutzt werden: Im Prinzip kann gesagt werden, dass alle Dotierkonzentrationen mit Werten oberhalb einer horizontalen Linie (dem Arbeitsabstand) detektierbar sind.

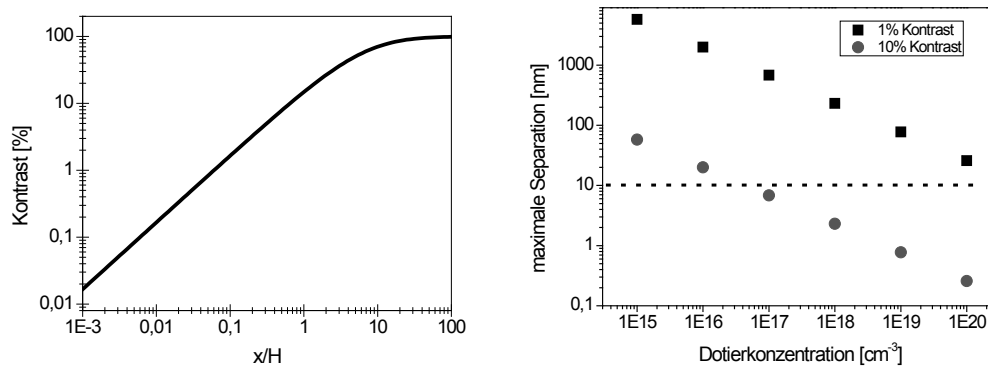


Abbildung 3.30: Auswertung des Einflusses der Separation auf den Kontrast. Offenbar gibt es eine universelle Kurve, die den zu erwartenden Kontrast in Abhängigkeit der relativen Weite der Verarmungszone abbildet (links). Aus dieser Darstellung lässt sich ein Arbeitsdiagramm (rechts) erzeugen.

Eine weitere, an dieser Stelle nicht diskutierte, Ursache für die Abnahme des Kontrastes ist, dass der minimale Wert der Krümmungskurve nicht zugänglich ist, da die Schwellspannung stark mit dem Abstand zwischen Spitze und Probe zunimmt. Man kann also lediglich einen Wert in der Nähe der beschriebenen Zustände bestimmen, was aufgrund der unterschiedlichen Krümmungsspannungs-Abhängigkeit zusätzlich den Kontrast verschlechtert.

Eine äußerst interessante Messmethode, bisher nicht in der Literatur beachtet, ergibt sich, wenn man die nächsthöhere Ableitung der Kraft-Spannungskurven untersucht. Eine Anwendung der Ableitung auf (3.4.3) ergibt:

$$\begin{aligned} \frac{\partial^3 F_{Tip}}{\partial V_T^3} &= \frac{1}{\varepsilon_0 \cdot A} \cdot \left(\frac{\partial Q_{Tip}}{\partial V_T} \cdot \frac{\partial^2 Q_{Tip}}{\partial V_T^2} + Q_{Tip} \cdot \frac{\partial^3 Q_{Tip}}{\partial V_T^3} + 2 \cdot \frac{\partial Q_{Tip}}{\partial V_T} \cdot \frac{\partial^2 Q_{Tip}}{\partial V_T^2} \right) \\ &= \frac{1}{\varepsilon_0 \cdot A} \cdot \left(3 \cdot \frac{\partial Q_{Tip}}{\partial V_T} \cdot \frac{\partial^2 Q_{Tip}}{\partial V_T^2} + Q_{Tip} \cdot \frac{\partial^3 Q_{Tip}}{\partial V_T^3} \right) = \frac{1}{\varepsilon_0 \cdot A} \cdot (3 \cdot C \cdot C' + Q_{Tip} \cdot C'') \end{aligned} \quad (3.4.8)$$

Zu erwarten ist also ein Signal, das vergleichbar zur Ableitung der differentiellen Kapazität eines MOS-Systems nach der Spannung ist. Da die Messung von dC/dV_T den Einfluss der

spannungsunabhängigen Kapazität der Lücke zwischen Spitze und Probe minimiert, ist ein hoher Kontrast zu erwarten.

Die grafische Darstellung (Abbildung 3.31 links) zeigt den erwarteten Verlauf: in der Nähe der Flachbandspannung (hier bei etwa -1V) zeigt sich sowohl in der Ableitung der Kapazität als auch in der dritten Ableitung der Kraft ein deutlich ausgeprägter Peak. Der zweite Peak (hier bei etwa 2V) erklärt sich durch den Wiederanstieg der Kapazitäts-Spannungs-Kurve beziehungsweise der Krümmungs-Spannungs-Kurve am Übergang zur Inversion. Da die Schwellspannung, die für diesen Übergang notwendig ist, wie bereits besprochen recht hoch sein kann, wird dieses zweite Signal im Folgenden nicht genauer betrachtet.

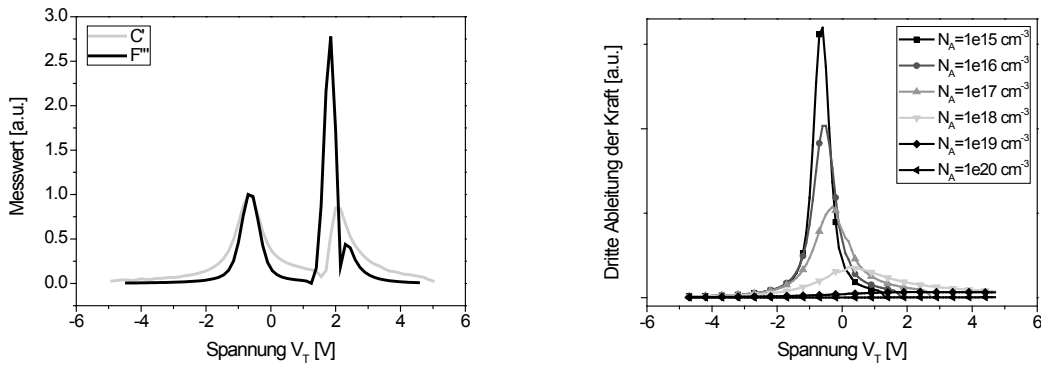


Abbildung 3.31: Vergleich zwischen dC/dV_T und d^3F/dV_T^3 für eine Separation von 5 nm (links) und Verlauf von d^3F/dV_T^3 für verschiedene Dotierkonzentrationen (rechts). Die starke Variation in der Höhe des Peaks lässt große Auflösung erwarten.

Die Höhe des Peaks hängt stark von der Dotierkonzentration ab, was leicht aus Gleichung (3.4.8) ersichtlich wird. Zur Abschätzung kann man annehmen, dass sich die mittlere Kapazität wenig ändert und der Wert der dritten Ableitung der Kraft tatsächlich proportional zur Ableitung der Kapazität ist. Diese kann nun wiederum wie folgt genähert werden:

$$\frac{dC}{dU} \approx \frac{C_{\max} - C_{\min}}{V_T - V_{FB}} = \frac{C_{TS} - \frac{1}{\frac{1}{C_{TS}} + \frac{x}{A_{TS} \cdot \epsilon_0 \cdot \epsilon_{Si}}}}{2 \cdot \phi_F + \frac{\sqrt{4 \cdot \epsilon_0 \cdot \epsilon_{Si} \cdot N \cdot \phi_F}}{C_{TS}}} \quad (3.4.9)$$

Nach einigen Umformungen ergibt somit ein Verlauf der Form:

$$\frac{\partial^3 F}{\partial U^3} \sim \frac{A}{B + C\sqrt{N}} \quad (3.4.10)$$

Für (3.4.10) wurden die Terme rigoros abgeschätzt, die Gleichung stellt also eine wirklich grobe Abschätzung dar. Stellt man allerdings den Wert der dritten Ableitung der Kraft zusammen mit Ausgleichskurven der genannten Form dar, ergibt sich doch recht gute Übereinstimmung (Abbildung 3.32a). Offensichtlicher wird das generelle Verhalten bei Auftragung in Koordinaten normiert auf die Werte bei $N_A = 10^{15} \text{ cm}^{-3}$. Alle Datenpunkte liegen in der Nähe eines simplen

Potenzgesetzes der Form $F(N_A) \propto F(10^{15}) \cdot (N_A/10^{15})^{-0.6}$, was erfreulicherweise einfache qualitative Aussagen erlaubt (Abbildung 3.32b).

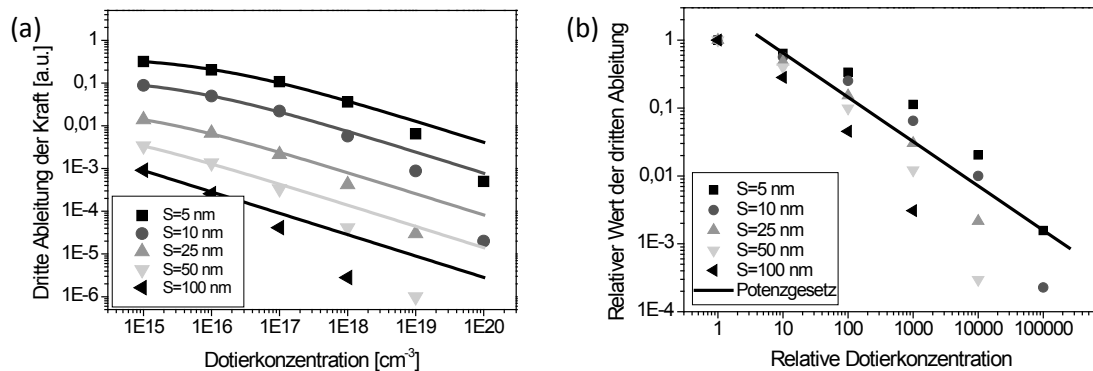


Abbildung 3.32: Abhängigkeit der dritten Ableitung der Kraft von der Dotierkonzentration für verschiedenen Separationen (a). Die Darstellung in relativen Koordinaten verdeutlicht die Abhängigkeit gemäß einem einfachen Potenzgesetz (b).

Zusammenfassend kann also formuliert werden:

1. Alle dem Experiment zugänglichen harmonischen Komponenten der Kraft zeigen zusätzlich zu ihrer Separationsabhängigkeit eine Abhängigkeit von der Dotierung der Probe.
2. Mit der SKM-Methode wird offenbar der Wert der Austrittsarbeit abgebildet, was einen Kontrast proportional zum Logarithmus der Dotierdichte ergibt.
3. Die SCM-Methode ähnelt Kapazitätsmessungen in MOS-Systemen. Definiert man als Messwert das minimale Signal, ergibt sich ein Kontrast etwa proportional zur Dotierdichte.
4. Die Abbildung der dritten Ableitung der Kraft ähnelt dem Verlauf nach der Ableitung der Kapazität eines MOS-Systems nach der Spannung, der maximal messbare Wert verhält sich indirekt proportional zur Wurzel der Dotierdichte.

3.5 Einfluss von Grenzflächenladungen auf die Messwerte

Auch im Fall von halbleitenden Proben stellt das Vorhandensein von fixierten Ladungen an der Grenzfläche zwischen der Probe und Vakuum, beziehungsweise innerhalb einer dielektrischen Schicht auf der Oberfläche, einen wichtigen Einfluss dar. Um diesen numerisch auszuwerten, wird die Randbedingung entsprechend angepasst (Rand 3 in Abbildung 3.24):

$$D_{\text{Vakuum,normal}} - D_{\text{Halbleiter,normal}} = \sigma \quad (3.5.1)$$

Die Variablen $D_{X,\text{normal}}$ stehen dabei für die Normalkomponenten der dielektrischen Verschiebung in Probe und Vakuum, σ ist wiederum die Flächenladungsdichte. Der Fall

geladener Schichten soll an dieser Stelle nicht behandelt werden, da gemäß den Ausführungen in 3.3 effektive Parameter wirksam sein sollten.

Aus allgemeiner Erfahrung bei Messungen von Halbleiterstrukturen [23] erwarten man zwei Auswirkungen von Ladungen: Zum Einen eine Änderung des Oberflächenpotentials und zum Anderen eine Verschiebung entlang der Spannungsachse der kapazitiven Signale. Beides kann durch Messung der Kraftwirkung abgebildet werden, wie Abbildung 3.33 illustriert. Dazu wurde die Abhängigkeit der entsprechenden Kraftkomponenten für ein typisches System in Abhängigkeit der Ladungsdichte simuliert. Offensichtlich ergeben sich die besprochenen Auswirkungen: Eine Verschiebung des Minimums der ersten Ableitung der Kraft und eine Verschiebung der Spannungsachse der zweiten Ableitung.

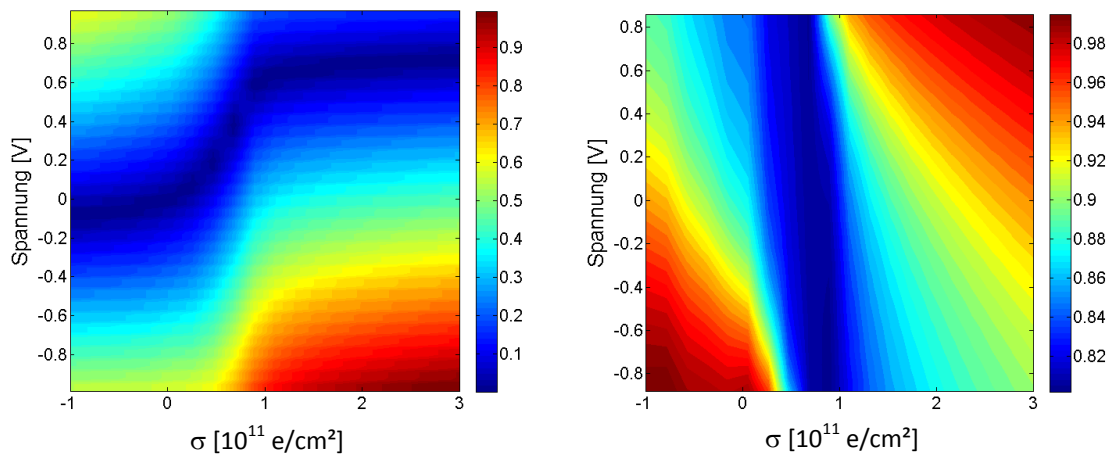


Abbildung 3.33: Darstellung der auf die Maxima normierten Komponenten der Kraftmessung in Abhängigkeit einer Grenzflächenladung σ für halbleitende Proben. Während sich das Minimum der ersten Ableitung (links) zu höheren Spannungen verschiebt, verschieben sich die Kurven (Spalten des Bildes) in der zweiten Ableitung (rechtes Bild) zu niedrigeren Spannungen.

Die absoluten Werte dieser Verschiebung sind unterschiedlich. Während in der Abbildung der ersten Ableitung eine nichtlineare Änderung der Position in Richtung positiver Spannung zu erkennen ist, verschieben sich die kapazitiven Signale (zweite Ableitung) offenbar linear in Richtung negativer Spannungen. Des Weiteren ändert sich der Verlauf der zweiten Ableitung derart, dass das Material invertiert erscheint: Während ohne zusätzliche Grenzflächenladung das Minimum der Kurven bei positiver Spannung zu finden ist, ist es bei positiven Ladungsdichten im negativen Bereich.

Erklären lassen sich die Effekte wie folgt: Eine Grenzflächenladung kann zum Beispiel in Gleichung (A.10) als vorgegebener Wert des elektrischen Feldes an der Oberfläche eingesetzt werden. Dadurch erhält man, da statt des Potentials wie dort beschrieben das Feld gefordert wird, die Umkehrkurven der Abbildung A.2, was dem Verlauf des Minimums in Abbildung 3.33a entspricht. Bei den kapazitiven Signalen ergibt sich die Verschiebung gemäß (3.5.2) durch Anwendung eines einfachen Plattenkondensatormodells:

$$\Delta V = -\frac{\sigma A_{TS}}{C_{TS}} \quad (3.5.2)$$

Grundlegende Wechselwirkungen im System Spitze-Probe

Das „Kippen“ der Kurven ist wiederum durch die Änderung des Oberflächenpotentials gegeben: Alle Messungen kapazitiver Eigenschaften bestimmen prinzipiell Änderungen der Ladungsträgerdichten in der Probe in der Nähe des Zustandes ohne äußere Spannung. Da eine Verschiebung des Oberflächenpotentials auch eine Verschiebung des Ferminiveaus bedeutet, verhält sich das System ab einem bestimmten Wert der Grenzflächenladungsdichte wie ein Halbleiter anderer Dotierung.

Es ist zu erwarten, dass die Dotierkonzentration Einfluss auf die beobachteten Effekte hat. Nutzt man die Verarmungsnäherung zur Beschreibung der Änderung des Oberflächenpotentials für kleine Werte der Ladungsdichte σ , ergibt sich inverse Proportionalität gemäß (3.5.3). Zur Herleitung wurde dazu die Poissongleichung unter Beachtung der Randbedingung gelöst.

$$\sigma = qN_A x$$

$$\varphi = \frac{\sigma^2}{2qN_A \epsilon_0 \epsilon_r} \quad (3.5.3)$$

Stellt man die Kurven $U_{min}(\sigma, N_A)$ der Spannungen an den minimalen Werten der ersten Ableitung dar, also dem Wert, welchen man aus SKM-Messungen erhalten sollte, ergibt sich Abbildung 3.34. Die Spannungen sind dabei jeweils in Bezug auf die Flachbandspannung (V_{FB}) angegeben, weshalb alle Kurven durch den Ursprung gehen. Die dargestellten Werte decken sich mit der Vorhersage, dass der Einfluss einer Oberflächenladung mit der Dichte der Dotierstoffe abnimmt. Außerdem ist zu erkennen, dass für größere Werte der Dotierkonzentration das einfache Modell der Verarmung besser anwendbar ist.

Offenbar ist, wie im Fall metallischer Proben, auch im Fall halbleitender Proben der Ladungszustand durch Kraftmessung zugänglich. Die quantitative Auswertung ist allerdings aufgrund der Komplexität der beschreibenden Gleichungen schwieriger. Auch ist zu erwarten, dass die Empfindlichkeit verglichen mit herkömmlichen kapazitiven Methoden höher ist, da diese zur Evaluierung eine Formel der Form (3.5.2) nutzen, was im Falle großer Kapazitäten (dünne Schichten, high-k Materialien) äußerst geringe Messwerte bedeutet.

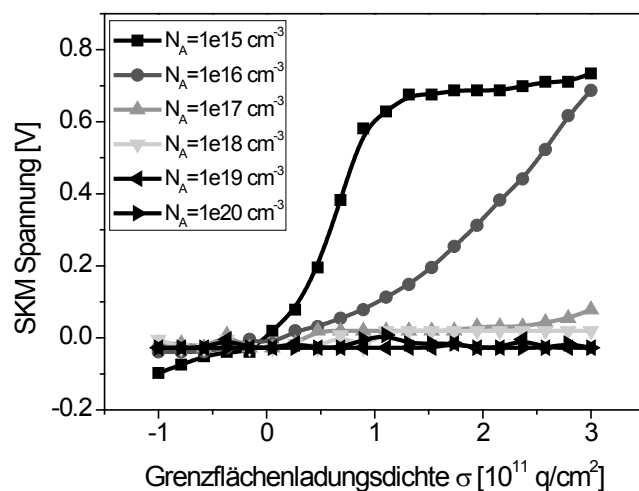


Abbildung 3.34: Simulierte Abhängigkeit der SKM-Spannung von der Grenzflächenladungsdichte für verschiedene Dotierkonzentrationen. Es ergibt sich, dass mit zunehmender Dichte der Dotieratome das Oberflächenpotential bei gleicher Ladungsdichte abnimmt.

Ausgehend von den hier vorgestellten Ergebnissen soll abschließend verallgemeinert werden:

1. Im Falle halbleitender Proben misst SKM das Oberflächenpotential. Auswirkungen darauf kann allerdings neben Ladungen auch Beleuchtung haben.
2. SCM misst kapazitive Eigenschaften eines halbleitenden Materials mit dem tatsächlich an der Oberfläche vorliegenden Fermi-niveau. Dieses Niveau bestimmt die Ladungsträgerdichten und unterliegt ebenfalls äußeren Einflüssen.

3.6 Die Auflösung im Fall halbleitender Proben

Einige Gedanken zur Auflösung der Methoden wurden für den Fall metallischer Proben bereits geäußert, allerdings ist zu erwarten, dass für halbleitende Proben zusätzliche Auswirkungen auftreten. So ist zum Beispiel davon auszugehen, dass laterale Effekte, wie zum Beispiel Übergänge in der Dotierkonzentration deutlichen Einfluss auf alle Messungen haben. Eine Modellierung solcher Experimente sollte außerdem wertvolle Informationen liefern, die aus der bisher lediglich quasi eindimensional eingeführten Theorie nicht erhältlich sind.

Aus Gründen der Komplexität werden einige Vereinfachungen vorgenommen:

1. Es werden einfache Systeme besprochen. An dieser Stelle soll das stellvertretend ein eindimensionaler p-n-Übergang sein.
2. Da aus Gründen der Symmetrie drei Dimensionen zur Simulation gewählt werden müssen, soll die Spitze des AFM-Systems auf den Apex reduziert werden. Wie in den Überlegungen zuvor gezeigt, scheint der Apex tatsächlich das auflösungsbestimmende Element zu sein, was die Einschränkung rechtfertigt.

Eine Skizze der Modellgeometrie zeigt Abbildung 3.35: Die als Kugel vereinfachte Spitze wird in einer gewissen Höhe über einen p-n-Übergang geführt. Durch Variation der Spannung an dieser Kugel werden folgend die entsprechenden Parameter, wie bereits beschrieben, extrahiert.

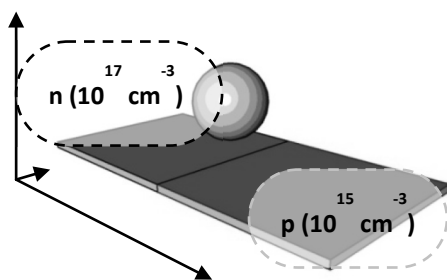


Abbildung 3.35: Vereinfachtes 3D-Setup für die Simulation der Abbildung eines p-n-Übergangs. Die an der als Kugel modellierten Spitze angelegte Spannung wird variiert, was Zugang zu den zu erwartenden Messgrößen bietet.

Als erstes sollen die Ergebnisse des virtuellen SKM-Experimentes besprochen werden. Dazu sind in Abbildung 3.36 die positionsabhängigen Signal-Spannungs-Kurven (erste Ableitung der Kraft) als Karte dargestellt (linkes Bild). Die Position des Übergangs ist dabei bei $x=0$ zu finden. Offenbar verläuft das Minimum des Signals entlang einer bestimmten Spannungs-Ort-Kurve, was

wie erwartet den Potentialverlauf am Übergang abbildet. Einige Spuren dieser Kurven für verschiedene Separationen sind ebenfalls angegeben, deutlich ist der Einfluss des Messabstandes zu bemerken (rechtes Bild). Mit zunehmender Separation verschmieren die Kurven maßgeblich, was wiederum ein Effekt der abstandsabhängigen Mittelung ist. Um nun quantitative Aussagen über die Auflösung zu erhalten, wurde folgender Ansatz gewählt: Unter Annahme einer ähnlichen Transferfunktion wie im Fall metallischer Proben kann davon ausgegangen werden, dass die modellierten Spuren durch Faltung des tatsächlichen Potentialverlaufs mit einer Lorentzfunktion entstehen. Mittels nichtlinearer Regression wurden nun die freien Parameter einer solchen Funktion bestimmt, die die Messwerte am besten erzeugt. Die Weite der so bestimmten Funktionen ist dann ein Maß für die Auflösung.

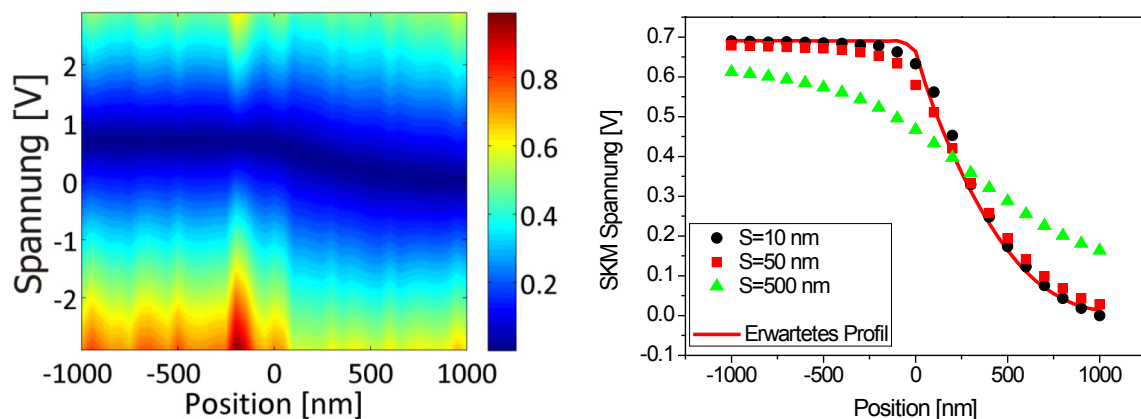


Abbildung 3.36: Karte der ersten Ableitung der elektrostatischen Kraft auf eine Kugel in Abhängigkeit der Spannung und der Position bezüglich eines p-n-Übergangs. Die Spuren des Minimums dieses Signals geben die erwartete SKM-Spannung an und sind für verschiedene Separationen im rechten Bild dargestellt. Offenbar gibt es mit zunehmender Höhe ein Glätten des gemessenen Profils, sodass der tatsächliche Verlauf (Linie) nicht mehr gemessen werden kann.

Die so erhaltenen Weiten der Transferfunktion sind in Abbildung 3.37 dargestellt. Augenscheinlich zeigt sich auch im Fall halbleitender Proben das bereits bekannte Verhalten der Weite der Transferfunktion: Für Separationen kleiner als der Radius des Apex steigt die Weite etwa proportional zur Wurzel des Radius, für größere Abstände proportional zum Radius.

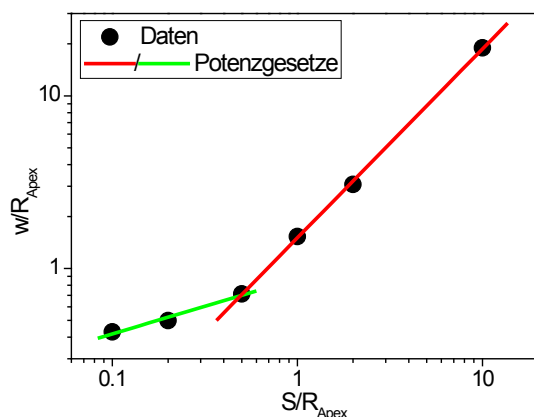


Abbildung 3.37: Durch rechnerische Entfaltung gewonnene Weiten der Transferfunktion für den kugelförmigen Apex im Falle des p-n-Übergangs. Der generelle Verlauf der Daten zeigt dieselbe Form wie für metallische Proben, man kann also Auflösung im Bereich des Apexradius auch für halbleitende Proben erwarten.

Für das gewählte Beispiel hat offenbar der Fakt der unterschiedlichen Kapazitäten der verschiedenen Bereiche zumindest geringen Einfluss auf die Messergebnisse, was eine Übernahme der Erläuterungen bezüglich des Auflösungsvermögens aus den früheren Kapiteln nahelegt.

Komplizierter, aber auch ungleich interessanter, ist die Auswertung der zu erwartenden kapazitiven Signale. Die zweite Ableitung der Kraft nach der Spannung, welche dem SCM-Signal gleicht, sollte nach den Aussagen der vorigen Abschnitte Informationen über die lokale Kapazität und somit über die lokale Dotierkonzentration ergeben. Die Frage nach dem Verhalten dieses Signals an Inhomogenitäten blieb bisher ungeklärt. Wie die Karte in Abbildung 3.38 (links) eindrucksvoll zeigt, ist auch im Fall von Übergängen ein starker Kontrast zu erwarten. Allerdings ist, wie der Verlauf des blau dargestellten Bereichs deutlich macht, ein Einfluss des Potentialverlaufs sichtbar. Die Messung erlaubt somit nicht eine direkte Abbildung des Dotierverlaufs, da lediglich Scans mit fester Spannung durchgeführt werden. Indes ergeben sich aber zwei mögliche Abbildungsmodi, wie die rechte Grafik in Abbildung 3.38 zeigt. Werden nämlich Profile mit den in der linken Grafik markierten Spannungen gemessen, ergibt sich wahlweise starker Kontrast am Übergang (schwarze Linie) oder, wie möglicherweise gewünscht, ein Dotierkonzentrationskontrast (rote Linie). Obwohl diese Ergebnisse bezüglich realer Experimente optimistisch stimmen, bleibt darauf hinzuweisen, dass die Signaländerungen in Bereichen vorliegen, die der (elektrostatischen) Breite des Übergangs entsprechen. Das bedeutet wiederum, dass für Strukturbreiten kleiner als die Debyelängen nur der Vergleich mit modellierten Werten quantitative Aussagen erlaubt.

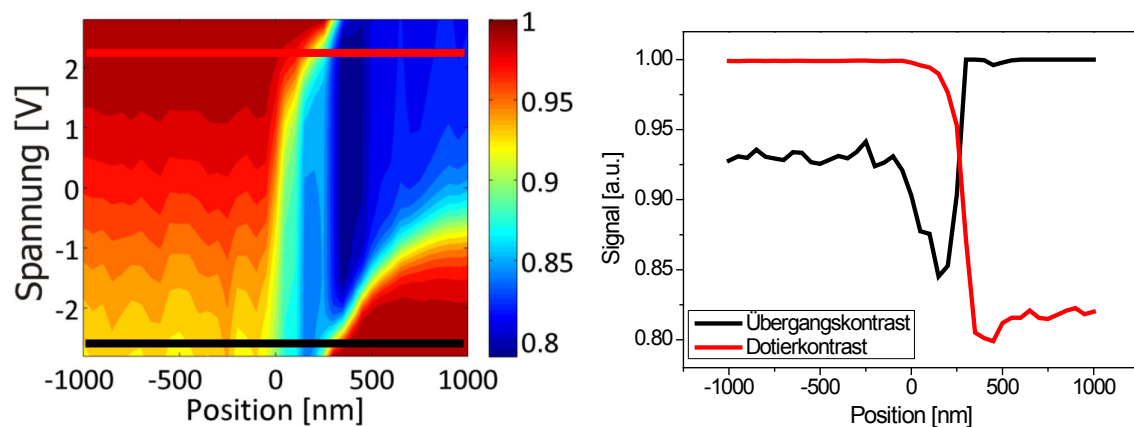


Abbildung 3.38: (Links) Karte eines simulierten SCM-Experiments am p-n-Übergang. Es ist starker, wenn auch spannungsabhängiger, Kontrast zu erkennen. Extrahiert man Linienprofile entlang der farblich markierten Spuren, ergeben sich unterschiedliche Abbildungsmodi (rechts). Je nach eingestellter Spannung kann so zwischen Lokalisierung des Übergangs und Abbildung des Dotierprofils gewählt werden.

Qualitativ ist zu erwarten, dass der Verlauf des SCM-Signals etwa dem Verlauf der Dichte freier Ladungsträger folgt. Die Potentialabhängigkeit dieses Wertes, was durch Verschiebung des lokalen Fermi-niveaus bedingt ist, erklärt den örtlichen Verlauf des kapazitiven Signals.

Wie zuvor erwähnt, kann die Auswertung der dritten Ableitung der Kraft nach der Spannung, also im Experiment die Kraftkomponente, welche mit der dritten Harmonischen der Anregung schwingt, wertvolle Informationen erbringen. Auch hierzu wurden für das Beispiel Karten

erzeugt, wie Abbildung 3.39 zeigt. Augenscheinlich kann man auch aus diesen Messungen deutlichen Kontrast erwarten, wie die Karte im linken Bild illustriert. Auch in diesem Fall ist es möglich, unterschiedliche Abbildungsmodi, je nach angelegter Spannung, einzustellen (rechtes Bild). Im Allgemeinen ist dieser zu erwartende Kontrast teilweise deutlich höher als in der zuvor beschriebenen Methode, die experimentelle Umsetzung kann somit ohne nennenswerten zusätzlichen Aufwand durchaus bessere Werte liefern.

Die zuvor gemachten Aussagen bezüglich des lokalen Signals gelten dabei auch für diese Methode, man misst also effektive Ladungsträgerdichten generiert durch die lokale Position des Fermi-niveaus. Da allerdings Ableitungen der Kapazität nach der Spannung maßgeblich das Signal bestimmen, ist sowohl der räumliche als auch parametrische Kontrast dabei höher abzuschätzen.

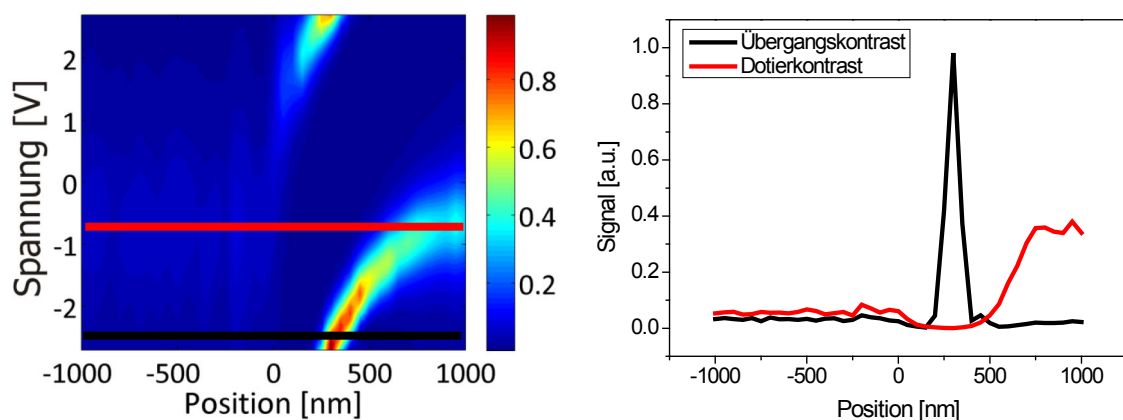


Abbildung 3.39: (Links) Karte eines simulierten 3ω -SCM-Experiments am p-n-Übergang. Wiederum ist ein starker spannungsabhängiger Kontrast messbar. Die extrahierten Linienprofile zeigen, dass auch in solchen Experimenten zwischen Lokalisierung des Übergangs und Abbildung des Dotierprofils gewählt werden kann. Allerdings ist der Kontrast verglichen mit der zuvor beschriebenen SCM-Methode stärker.

Ausgehend von den vorgestellten Ergebnissen kann also postuliert werden:

1. Die Auflösung der SKM-Methode bei halbleitenden Proben ist zumindest in derselben Größenordnung verglichen mit Experimenten an metallischen Proben. Die spannungsabhängige Krümmung der Signal-Spannungs-Kurve hat somit minimalen Einfluss auf die Ergebnisse und die für metallische Proben gewonnenen Aussagen können übernommen werden.
2. Die kapazitiven Signale erzeugen sowohl örtlichen als auch parametrischen Kontrast. Es kann dabei jeweils zwischen Abbildung des Übergangs und Abbildung eines Ladungsdichtenkontrastes gewählt werden.
3. Die Auflösung, beziehungsweise der Einflussbereich, der kapazitiven Messungen ist durch die Debyelängen limitiert. Innerhalb dieser Längenskale ändern sich die gemessenen Signale, eine deutlich genauere Lokalisierung von Störungen erscheint nicht möglich.

3.7 Methodischer Vergleich zwischen Kapazitäts- und Kraftmessung

Da üblicherweise Messungen von Halbleiterproben in Atomkraftmikroskopen durch Skalierung klassischer, großflächiger Messungen umgesetzt werden, sollen an dieser Stelle vergleichende Studien erläutert werden.

Tatsächlich ist es so, dass die Standard SCM-Methode die Kapazität des Spitze-Probe-Systems misst und Parameter von Interesse aus den Messwerten extrahiert. Gewöhnlich wird dazu die Spitze in Kontakt mit einer zusätzlich erzeugten Isolatorschicht auf der Oberfläche der Probe gebracht und mittels Standardmesstechnik ausgelesen. Da die Messungen im Kontakt-Modus durchgeführt werden, sind sowohl hohe räumliche Auflösung als auch minimale Rückkopplung zwischen Oberflächenabbildung und elektrischer Messung zu erwarten [24],[25].

Zur numerischen Nachstellung beider Methoden, also der Standardmessung und der in dieser Arbeit besprochenen Kraftmessung, wurden für eine Spitze mit typischen geometrischen Parametern sowohl die Ladung als auch die Kraft in Abhängigkeit von der angelegten Spannung bestimmt. Die unterschiedlichen experimentellen Umsetzungen wurden dabei durch unterschiedliche Separationen nachgestellt, hier 0.5 nm für die Ladungsmessung und 20 nm für die Kraftmessung.

Wie Abbildung 3.40 veranschaulicht, sind die extrahierten Verläufe von Kapazität und zweiter Ableitung der Kraft und auch von Ableitung der Kapazität und dritter Ableitung der Kraft sehr ähnlich. Allerdings ist ebenfalls offensichtlich ein Unterschied im erreichbaren Kontrast zu bemerken: So ändert sich zum Beispiel die Kapazität im Bereich von einem Prozent, das Kraftäquivalent dazu um zehn Prozent. Da diese Variation direkt mit der Sensitivität zu verknüpfen ist, erscheint die Kraftmessung günstiger. Erklärbar wird dieses Verhalten dadurch, dass immer Serienschaltungen von Proben- und Sondenkapazität gemessen werden, die kleinere von beiden bestimmt den Messwert. Im Fall der Kapazität C ist der Einfluss der Cantileverbasis sehr groß, demzufolge werden Variationen der Probenkapazität schlecht wiedergegeben. Im Gegensatz dazu ist der Einfluss der Basis für die Kraftmessung F deutlich vermindert, was eine sehr einfache Abschätzung zeigt:

$$\begin{aligned} C &= A_{Basis} \cdot c_{Basis} + A_{Apex} \cdot c_{Apex} \\ F &= \frac{1}{2} \cdot A_{Basis} \cdot \varepsilon \cdot E_{Basis}^2 + \frac{1}{2} \cdot A_{Apex} \cdot \varepsilon \cdot E_{Apex}^2 \\ \frac{\partial^2}{\partial V^2} F &= \frac{1}{\varepsilon} \cdot (A_{Basis} \cdot c_{Basis}^2 + A_{Apex} \cdot c_{Apex}^2) + g \left(E \cdot \frac{\partial c}{\partial V} \right) \end{aligned} \quad (3.7.1)$$

Wenn man von der in (3.7.1) genutzten, quasi eindimensionalen, Beschreibung ausgeht, erkennt man, dass für die Kraftmessung die Abstandsabhängigkeit der Flächenkapazitäten c eine größere Rolle spielt als für die Kapazität. Schätzt man weiter die Einträge ab ergibt sich:

$$\begin{aligned} C &\sim \frac{R_{Basis}^2}{H+S} + \frac{R_{Apex}^2}{S} \Rightarrow v_C = \frac{R_{Basis}^2}{R_{Apex}^2} \frac{S}{H+S} \approx 300 \\ \frac{\partial^2}{\partial V^2} F &= \frac{R_{Basis}^2}{(H+S)^2} + \frac{R_{Apex}^2}{S^2} \Rightarrow v_F = \frac{R_{Basis}^2}{R_{Apex}^2} \frac{S^2}{(H+S)^2} \approx 1 \end{aligned} \quad (3.7.2)$$

Die in (3.7.2) abgeschätzten Verhältnisse ν unterscheiden sich sehr stark, was den Unterschied in der Sensitivität erklärt. Die tatsächlichen Verhältnisse können dabei Abbildung 3.3 entnommen werden.

Die dritte Ableitung der Kraft nach der Spannung zeigt, wie auch Gleichung (3.4.8) zu entnehmen, etwa den gleichen Verlauf wie die Ableitung der Kapazität nach der Spannung. Da der Einfluss der Cantileverbasis in beiden Methoden unterdrückt wird, ist auch der Kontrast vergleichbar.

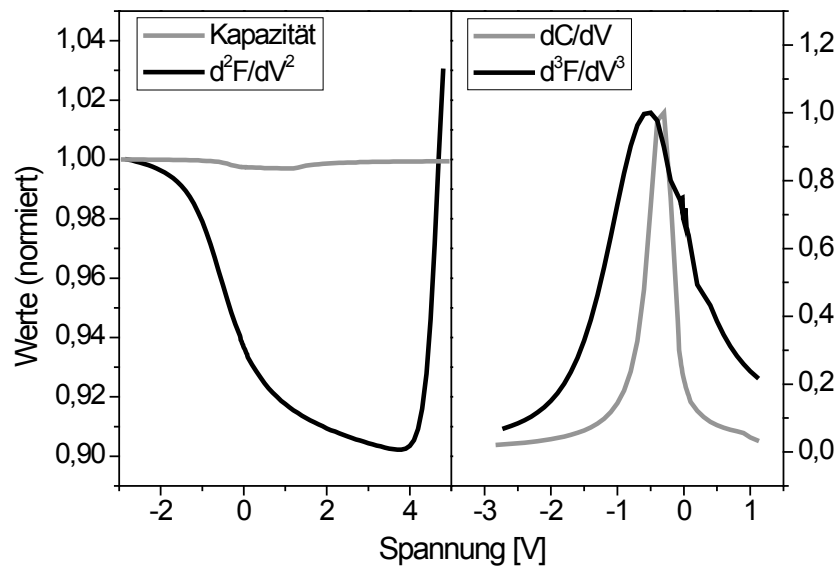


Abbildung 3.40: Vergleich von Signal-Spannungskurven für Kapazitäts- und Kraftmessungen. Obwohl die Signalverläufe sich ähneln, ist der zu erwartende Kontrast im Fall der Kraftmessung höher.

Nachdem prinzipielle Unterschiede und Übereinstimmungen abgeklärt sind, erscheint es wünschenswert, den parametrischen Kontrast zu behandeln. In diesem Fall wäre das die Abhängigkeit der Signale von der Dotierkonzentration der Probe. Die Ergebnisse der Auswertung zeigt Abbildung 3.41.

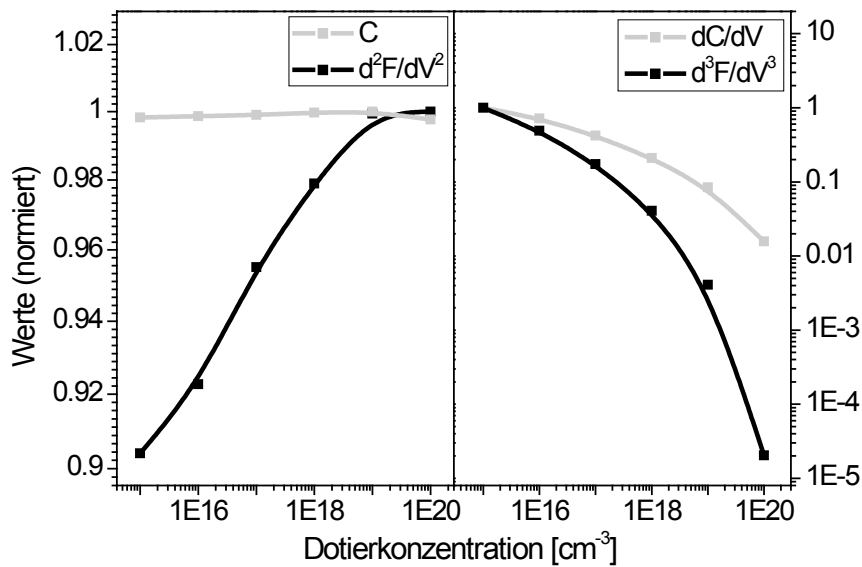


Abbildung 3.41: Sensitivität von Kapazität- und Kraftmessung. Während das kapazitive Signal keine Unterscheidung der Dotierung erlaubt, ist für die Kraftmessung ein messbarer Unterschied vorhanden (links). Die Ableitung der Kapazität unterscheidet sich indes wenig vom kraftbasierten Gegenstück, lediglich die Größenordnung des erreichbaren Kontrastes ist eine andere.

Der bereits besprochene Effekt der Unterdrückung des Einflusses der Cantileverbasis auf die Kraftmessung zeigt sich eindrucksvoll am Verlauf der Dotierabhängigkeit der Signale. Während bei Messung der Kapazität kein Unterschied zwischen hoch und niedrig dotierten Proben zu erwarten ist, zeigt die Auswertung des kraftbasierten Signals merklichen Kontrast. Dieser Unterschied ist bei Verwendung der Kapazitätsableitung beziehungsweise des Gegenstücks nicht so prominent, lediglich die zu erwartenden Größenordnungen unterscheiden sich.

Offenbar gilt zusammenfassend:

1. Kraftbasierte Methoden erlauben vergleichbare Messungen wie kapazitätsbasierte.
2. Der zu erwartende Dotierkontrast kann teilweise merklich variieren.

Beachtet man zusätzlich die Notwendigkeit einer zusätzlichen isolierenden Schicht für die Messung der Ladung, erscheinen die ohne tatsächlichen Kontakt zur Probe durchgeführten Kraftmessungen äußerst vielversprechend.

4. Die Transferfunktion

Wie bereits in den vorhergehenden Kapiteln erwähnt, ist das Wissen über die Transferfunktion von zentraler Bedeutung, wenn es um die Beurteilung von Auflösung und Sensitivität eines EFM-Experimentes geht. Obwohl diese Funktion gemäß ihrer Definition lediglich die Antwort eines SKM-Systems, also das gemessene Oberflächenpotential, angibt, erscheint es sinnvoll, die gefundenen Eigenschaften auch auf andere Methoden zu übertragen.

In diesem Kapitel soll versucht werden, sowohl analytische als auch numerische Aussagen über diese Transferfunktion zu gewinnen. Ein Vergleich der gewonnenen Abhängigkeiten sollte es ermöglichen, nützliche „Faustformeln“ für die experimentelle Arbeit zu erhalten.

4.1 Näherungsweise analytische Bestimmung der Transferfunktion

Um einen, zumindest näherungsweise gültigen, analytischen Ausdruck für die Transferfunktion zu erhalten, wird die Spitze des AFM-Systems in ihre drei signifikanten Teilbereiche zerlegt. Der hier gewählte Ansatz wurde durch Kapazitätsschätzungen in [26] motiviert, er kann allerdings zur Auflösungsbestimmung modifiziert werden.

4.1.1. Der kugelförmige Apex

Ein intuitiver Ansatz ist die Zerlegung der Oberfläche in Facetten, also eine stückweise lineare Näherung der kreisförmigen Begrenzung der Oberfläche im (r,z) -Koordinatensystem. Ausgehend von jedem Punkt dieser Begrenzungslinie kann dann ein Bogen zur Fläche $z=0$ geschlagen werden, dessen Länge zur Berechnung interessanter Größen herangezogen werden kann.

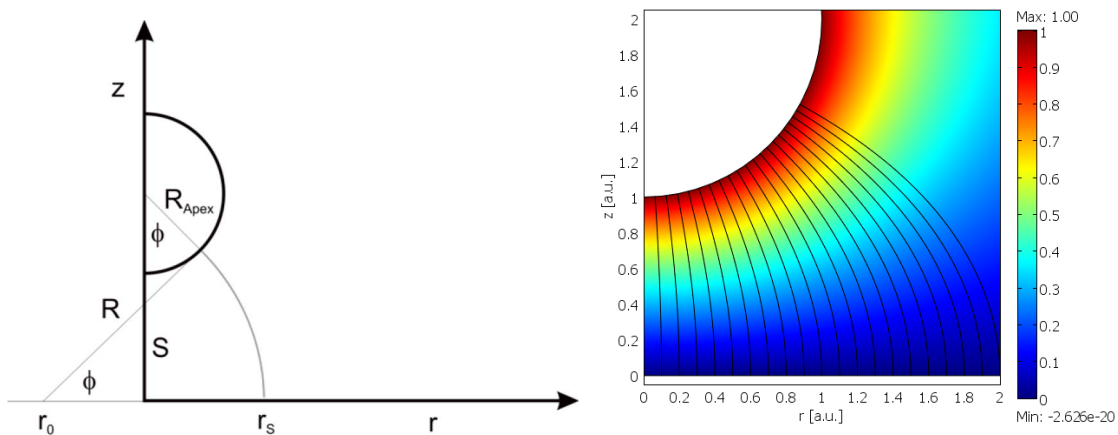


Abbildung 4.1: Skizze mit den für die Facettierung der Kugel benutzten Größen (links) und bildliche Darstellung der Feldlinien gewonnen aus einer numerischen Lösung des Problems (rechts). Offenbar ist die Annahme von Kreisbögen sinnvoll.

Gemäß der Skizze in Abbildung 4.1 ist folgendes Gleichungssystem zu lösen:

$$\begin{pmatrix} r_0 + R \cos(\varphi) \\ R \sin(\varphi) \end{pmatrix} = \begin{pmatrix} R_{Apex} \sin(\varphi) \\ S + R_{Apex} - R_{Apex} \cos(\varphi) \end{pmatrix} \quad (4.1.1)$$

Die Transferfunktion

Darin entsprechen r_0 dem Schnittpunkt der Tangente am betrachteten Punkt mit der z-Achse, R_{Apex} dem Radius des Apex, S dem Abstand zwischen Apex und der Ebene $z=0$ und R schließlich dem gesuchten Radius des Bogens.

Die Lösung ergibt sich zu:

$$\begin{aligned} r_0 &= -\frac{\cos(\varphi)S + (\cos(\varphi) - 1)R_{Apex}}{\sin(\varphi)} \\ R &= \frac{S - (\cos(\varphi) - 1)R_{Apex}}{\sin(\varphi)} \end{aligned} \quad (4.1.2)$$

Die Länge L des Bogens und der Aufpunkt r_S auf der z-Achse errechnen sich zu:

$$\begin{aligned} L &= R \cdot \varphi = \varphi \cdot \frac{S + (1 - \cos(\varphi))R_{Apex}}{\sin(\varphi)} \\ r_S &= r_0 + R = \frac{(1 - \cos(\varphi))(S + 2R_{Apex})}{\sin(\varphi)} \end{aligned} \quad (4.1.3)$$

Offenbar ist es günstig, im Weiteren relative Größen zu benutzen, was durch Normierung auf den Radius des Apex geschieht:

$$\begin{aligned} s &= \frac{S}{R_{Apex}} \\ l &= \frac{L}{R_{Apex}} = \varphi \cdot \frac{s + 1 - \cos(\varphi)}{\sin(\varphi)} \\ x &= \frac{r_S}{R_{Apex}} = \frac{(1 - \cos(\varphi))(s + 2)}{\sin(\varphi)} \end{aligned} \quad (4.1.4)$$

Um die Validität der so gewonnenen Größen zu überprüfen, soll zunächst die Kapazität des Systems Apex/Ebene berechnet werden. Da das elektrische Feld invers proportional zur Länge l ist und die Oberflächenladungsdichte wiederum proportional zu diesem Feld gilt:

$$C = \int_0^\pi 2\pi\epsilon_0 R_{Apex} \frac{\sin^2(\varphi)}{\varphi(s + 1 - \cos(\varphi))} d\varphi \quad (4.1.5)$$

Unglücklicherweise ist dieses Integral nicht analytisch lösbar. Allerdings können folgende Approximationen erfolgen: Eine Taylorentwicklung der Winkelfunktionen liefert eine untere Schranke gemäß:

$$\int_0^\pi \frac{\sin^2(\varphi)}{\varphi(s + 1 - \cos(\varphi))} d\varphi > \int_0^{\frac{\pi}{2}} \frac{\varphi}{\left(s + \frac{\varphi^2}{2}\right)} d\varphi = \ln\left(1 + \frac{\pi^2}{8} \cdot \frac{1}{s}\right) > \ln\left(1 + \frac{1}{s}\right) \quad (4.1.6)$$

Dabei wurde außerdem angenommen, dass der Integrand auf der unteren Halbkugel grösser ist als auf der oberen. Mit genau dieser Annahme und der Feststellung, dass $\sin(x) < x$ im betrachteten Intervall gilt, erhält man eine obere Schranke gemäß:

$$\int_0^\pi \frac{\sin^2(\varphi)}{\varphi(s+1-\cos(\varphi))} d\varphi < 2 \int_0^{\frac{\pi}{2}} \frac{\sin^2(\varphi)}{\varphi(s+1-\cos(\varphi))} d\varphi < 2 \int_0^{\frac{\pi}{2}} \frac{\sin(\varphi)}{(s+1-\cos(\varphi))} d\varphi = 2 \ln\left(1 + \frac{1}{s}\right) \quad (4.1.7)$$

Es erscheint somit sinnvoll, die Kapazität wie folgt anzugeben:

$$C = 2\pi\epsilon_0 R_{Apex} \cdot K \cdot \ln\left(1 + \frac{1}{s}\right); 1 \leq K \leq 2 \quad (4.1.8)$$

In obiger Gleichung (4.1.8) wurde eine Konstante K eingeführt, deren Wert in etwa 1 beträgt. Zur Validierung kann nun eine analytische Beziehung für die Kapazität einer Kugel benutzt werden, die durch die Methode der Bildladungen bestimmt wurde [27]:

$$C = 4\pi\epsilon_0 R \cdot \left(1 + \sum_{n=2}^{\infty} \frac{\sinh(\alpha)}{\sinh(n\alpha)}\right); \cosh(\alpha) = 1 + s \quad (4.1.9)$$

Unter Beachtung der Tatsache, dass in der analytischen Beziehung zusätzlich die Eigenkapazität der Kugel $C=4\pi\epsilon_0 R$ enthalten ist, müssen zur Bestimmung der Konstanten folgende Funktionen verglichen werden:

$$f_A = 2 \cdot \sum_{n=2}^{\infty} \frac{\sinh(\alpha)}{\sinh(n\alpha)}, \cosh(\alpha) = 1 + s \quad (4.1.10)$$

$$f_B = K \cdot \ln\left(1 + \frac{1}{s}\right)$$

Das Ergebnis dieses Vergleichs für $K=1$ ist in Abbildung 4.2 dargestellt.

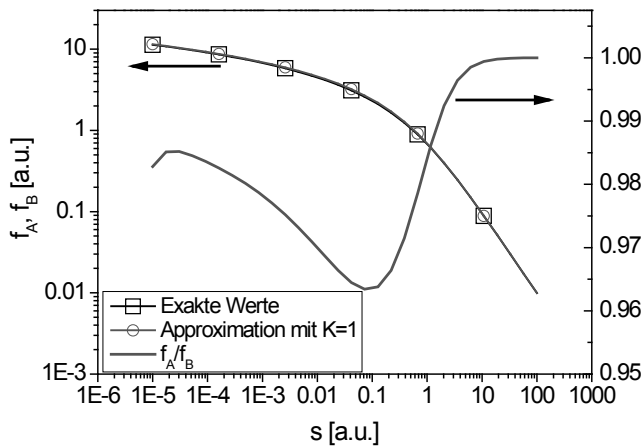


Abbildung 4.2: Vergleich der Ausdrücke für die Kapazität eines Kugel/Ebene-Systems. Die Werte für den exakten Ausdruck (f_A) und den hier gewonnenen (f_B) unterscheiden sich kaum, was die Nutzbarkeit der Theorie belegt.

Die Transferfunktion

Weil die Graphen der generierten Funktionen nur sehr wenig voneinander abweichen, wurde zusätzlich das Verhältnis der Funktionen dargestellt. Diese Größe erlaubt gleichzeitig eine Abschätzung des Fehlers, welcher damit im experimentell relevanten Bereich (Apexradien im Bereich von 50 nm, damit $s=S/R$ zwischen 10^{-2} und 10) maximal etwa 5% beträgt.

Da offenbar die mit der Facettierungsmethode erhaltenen Ausdrücke, für $K=1$, eine gute Aussage über die Kapazität des Systems erlauben, kann davon ausgegangen werden, dass auch die Positionsabhängigkeiten sinnvolle Näherungen darstellen. Um nun, wie erwünscht, eine geschlossene Form für die Transferfunktion anzugeben, wird ein Ausdruck für die wirkende Kraft in Abhängigkeit von der Koordinate r benötigt. Ausgehend von den über den Winkel φ parametrisierten Größen in Gleichung (4.1.4) erhält man, unter der Annahme, dass die Kraft proportional zum Quadrat des elektrischen Feldes ist, eine Vorschrift für die gewünschte Funktion. Da ein direkter Zusammenhang zwischen r -Koordinate und Kraft erhalten werden soll und ein einfaches Umstellen nicht analytisch durchführbar ist, werden die Größen x und l für kleine Winkel φ entwickelt. Man erhält:

$$\begin{aligned} x &\approx \frac{\varphi}{2} \cdot (s+2) \\ l &\approx s + \frac{(s+3)\varphi^2}{6} \\ F(x) &\sim \frac{1}{l^2(x)} = \frac{1}{\left(s + \frac{2(s+3)}{3(s+2)^2} x^2 \right)^2} \end{aligned} \quad (4.1.11)$$

Wie den Darstellungen der exakten und approximierten Kraft-Radius-Kurven (Abbildung 4.3) entnommen werden kann, sind die Abweichungen sowohl für kleine Abstände (Abbildung 4.3a, $s=0.1$) als auch für große Abstände (Abbildung 4.3b, $s=10$) marginal. Es erscheint somit gerechtfertigt, die Kraft-Radius-Abhängigkeit, die in Gleichung (4.1.11) eingeführt wurde, als nutzbar anzusehen.

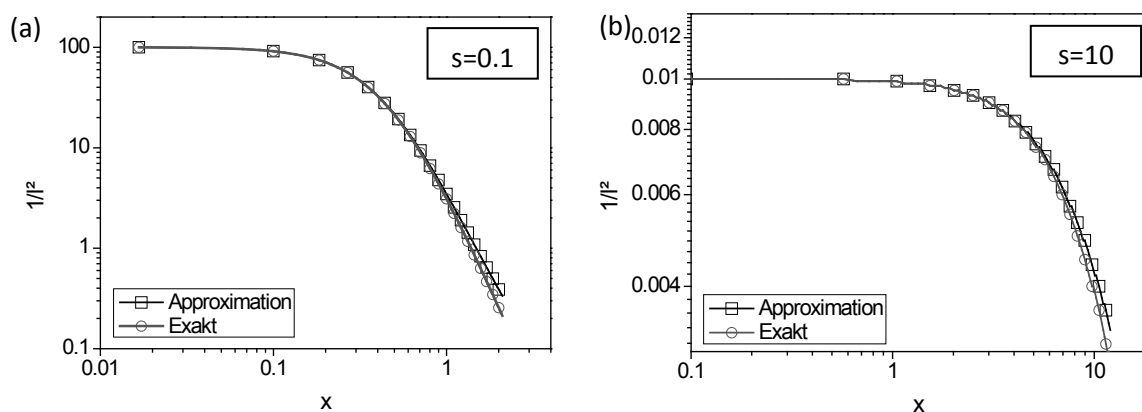


Abbildung 4.3: Vergleich der exakten und genäherten Ausdrücke für $1/l^2$ gemäß (4.1.11) und (4.1.4). Die Unterschiede sind sowohl für kleine Spitze-Probe-Abstände (a, $s=0.1$) als auch für große (b, $s=10$) vernachlässigbar.

Einige algebraische Umformungen führen zu einer gefälligeren Form für $F(x)$:

$$F(x) = \frac{F(0) \cdot w^4}{(w^2 + x^2)^2}; F(0) = \frac{C}{s^2}, w^2 = \frac{3s(s+2)^2}{2(s+3)} \quad (4.1.12)$$

Die Formulierung in (4.1.12) erlaubt ein einfaches Anpassen experimenteller Daten und gleichzeitig einfaches analytisches Abschätzen typischer Größen wie zum Beispiel des Einflussbereichs. Es gilt nämlich, dass die Kraft für $x=w$ auf 1/16 des Wertes für $x=0$ abgefallen ist. Die hier willkürlich eingeführte Größe w hat demnach tatsächlich die Bedeutung einer Breite.

Um Plausibilität der in (4.1.12) angegebenen Formeln zu überprüfen, wurde das Randwertproblem numerisch gelöst und das Quadrat des elektrischen Feldes entlang der r -Achse extrahiert. Da diese Größe proportional zur Kraft ist, muss eine Kurvenanpassung an eine Funktion der gewünschten Form möglich sein. Die Ergebnisse dieser Prozedur sind in Abbildung 4.4 dargestellt.

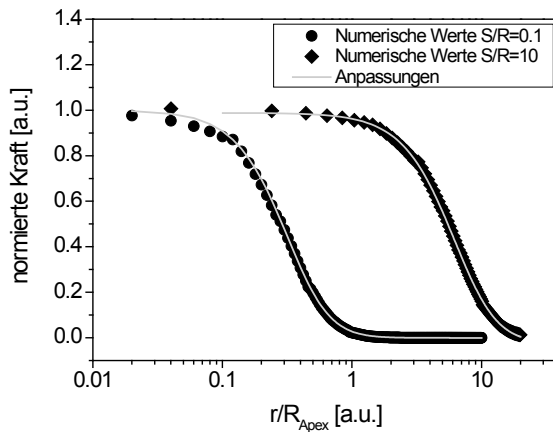


Abbildung 4.4: Kurvenanpassungen gemäß (4.1.12) an numerisch gewonnene Daten. Die Übereinstimmung der Datenpunkte mit den vorhergesagten Verläufen ist sowohl für kleine als auch große Abstände gut.

Wie der Grafik (Abbildung 4.4) entnehmbar ist, stimmen die extrahierten Kurven mit den Erwartungen überein. Eine weitere Verwendung der benutzen Näherung erscheint dadurch überaus sinnvoll. Besonders zweckmäßig ist es, die Abhängigkeit der Kurvenparameter ($F(0)$ und w) vom Abstand zwischen Kugel und Fläche zu untersuchen. Dazu wurden numerische Lösungen für Abstände im gewünschten Bereich ermittelt und die erwähnten Anpassungen durchgeführt. Die in Abbildung 4.5 dargestellten Ergebnisse dieses Vorgehens unterstreichen die Güte der analytischen Kraft-Abstandsabhängigkeit, da die errechneten Werte für die Weite w (Abbildung 4.5a) sehr gut zu den vorhergesagten passen. Interessanterweise lassen sich einfache Potenzfunktionen für w und $F(0)$ angeben, welche ebenfalls in Abbildung 4.5 gezeigt sind.

Diese Potenzfunktionen sind:

$$w = \begin{cases} \sqrt{2s} & s < 1 \\ s & s > 1 \end{cases} \quad (4.1.13)$$

$$F(0) = \begin{cases} \frac{a}{s^2} & s < 1 \\ \frac{b}{s^4} & s > 1 \end{cases}$$

Die Transferfunktion

Die Abhängigkeiten für große s stimmen im Übrigen mit denen überein, die auch aus einer Multipolentwicklung (vergleiche Anhang D) entnommen werden können.

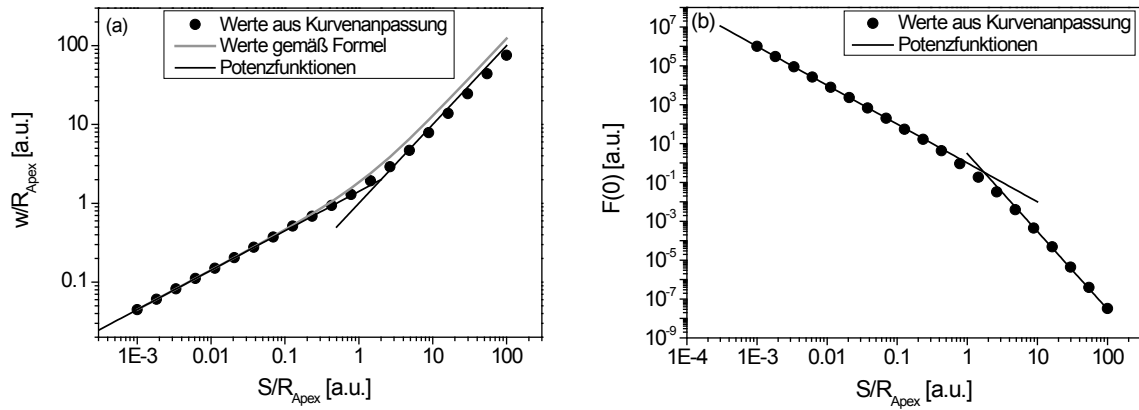


Abbildung 4.5: Ergebnisse der numerischen Überprüfung der in Gleichung (4.1.12) angegebenen Abhängigkeiten. Offenbar stimmt die extrahierte Breite der angepassten Funktionen (a, Punkte) gut mit den Vorhersagen (a, graue Spur) überein. Die bestimmte Maximalkraft (b, Punkte) folgt nur für Abstände kleiner als R_{Apex} der erwarteten Spur, lässt sich allerdings auch darüber hinaus einfach beschreiben.

Zur Bestimmung der Transferfunktion ist weiterhin notwendig, den in Richtung der z -Achse wirkenden Anteil der Kraft zu bestimmen. Offenbar gilt (vgl. Abbildung 4.1):

$$F_z(\varphi) = F(\varphi) \cdot \cos(\varphi) \quad (4.1.14)$$

Da dieses Ergebnis allerdings die Nutzung der parametrisierten Formulierung (siehe (4.1.4)) vorschreibt, soll auch eine Näherung der Form (4.1.12) gefunden werden. Dazu kann eine effektive Länge l_{eff} definiert werden, die dann entsprechend entwickelt wird:

$$l_{\text{eff}} = \frac{l}{\sqrt{\cos(\varphi)}} = \varphi \cdot \frac{s+1-\cos(\varphi)}{\sin(\varphi)\sqrt{\cos(\varphi)}} \quad (4.1.15)$$

$$l_{\text{eff}} \approx s + \frac{(5s+6)\varphi^2}{12} = s + \frac{(5s+6)x^2}{3(s+2)^2}$$

Ein Einsetzen liefert eine ähnliche Kraft-Radius-Abhängigkeit wie in Gleichung (4.1.12), allerdings ist die Weite w anders definiert:

$$w^2 = \frac{3s(s+2)^2}{5s+6} \quad (4.1.16)$$

Eine Abbildung der aus der Parametrisierung gewonnenen Kurven und der gefundenen Näherung für $s=0.1$ und $s=10$ entnimmt man Abbildung 4.6. Offenbar ist die Genauigkeit der Näherung für große Abstände nicht sehr hoch, allerdings kann eine Kurvenanpassung der Form (4.1.12) vorgenommen werden, was eine Nutzung dieser Formulierung auch für die z -Komponente der Kraft motiviert.

Die Bestimmung der Transferfunktion geschieht nun durch Normierung auf das Integral der Kraft:

$$T(x) = \frac{F_z(x)}{\int_0^{X_{Max}} 2\pi x F_z(x) dx} \approx \frac{F_z(x)}{\int_0^{\infty} 2\pi x F_z(x) dx} = \frac{w^2}{\pi(w^2 + x^2)^2} \quad (4.1.17)$$

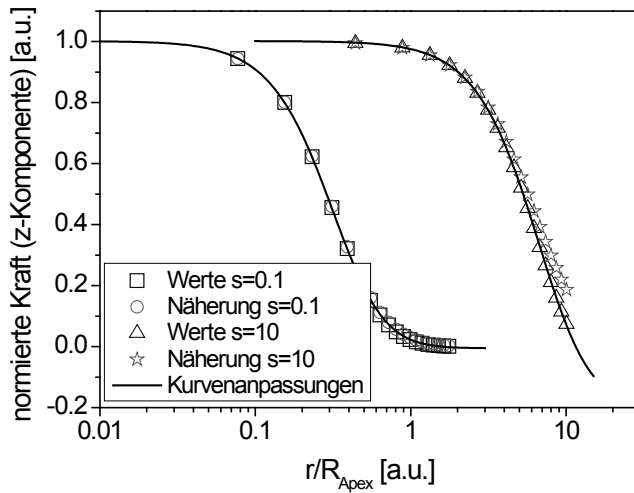


Abbildung 4.6: Verlauf der Normalkomponente der Kraft in Abhängigkeit der lateralen Koordinate. Die hergeleiteten analytischen Ausdrücke und deren Näherungen stimmen gut überein, quasi unabhängig von der Wahl der normierten Separation s .

Die in (4.1.17) angewendete Näherung entspricht einer Ausdehnung des Integrals über den im Rahmen des Modells gültigen maximalen Wert für x , X_{Max} , bestimmbar durch (4.1.4), bis ins Unendliche. Der so erzeugte Fehler ist relativ gering und kann analytisch abgeschätzt werden. Es stellt sich heraus, dass (4.1.17) lediglich geringfügig modifiziert werden muss:

$$T(x) = \frac{C(s)w^2}{\pi(w^2 + x^2)^2} \quad (4.1.18)$$

Die Funktion $C(s)$ nimmt Werte zwischen 2 und 2.5 an, kann also durchaus als konstant angenommen werden.

4.1.2. Der kegelförmige Teil

Wiederum unter der Annahme, dass das an einem Punkt r wirkende elektrische Feld E indirekt proportional zur Länge des Bogens zwischen diesem Punkt und einem korrespondierenden Punkt auf der Oberfläche des Kegels ist, ergibt sich:

$$E(r) = \frac{U}{R \cdot \varphi} = \frac{U}{\left(r - R_0 + \frac{S}{\tan(\varphi)} \right) \cdot \varphi} \quad (4.1.19)$$

Die Transferfunktion

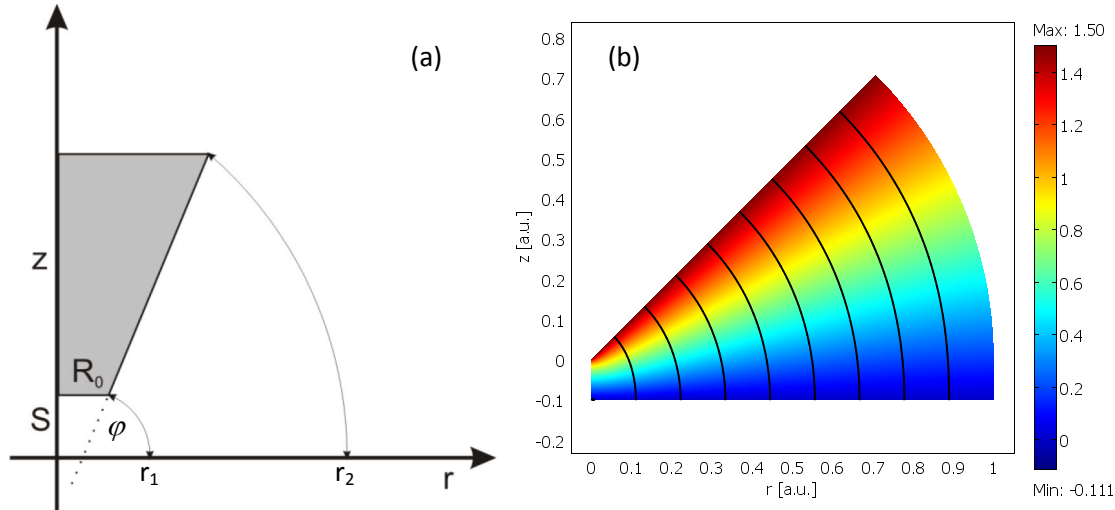


Abbildung 4.7: Verwendete Geometrie (a) und numerische Lösung des Randwertproblems (b). Dargestellt sind neben dem Verlauf des Potentials (Farbverlauf) auch einige Feldlinien des elektrischen Feldes (Linien)

In Gleichung (4.1.19) bedeutet U , die für den Moment als konstant angenommene, Spannung zwischen den betrachteten Punkten, R_0 den Radius des abgeschnittenen Teils des Kegels, S den Abstand zwischen Kegel und Oberfläche sowie φ den eingeschlossenen Winkel. Die entsprechende Kraft ergibt sich damit zu:

$$F(r) = \frac{1}{2} \varepsilon_0 E(r)^2 = \frac{1}{2} \varepsilon_0 U^2 \frac{1}{\left(r - R_0 + \frac{S}{\tan(\varphi)} \right)^2 \cdot \varphi^2} \quad (4.1.20)$$

Beachtet man weiterhin, dass die vom System gemessene Kraft lediglich dem Anteil der Gesamtkraft parallel zur z-Achse entspricht, ergibt sich schließlich:

$$F_z(r) = \frac{1}{2} \varepsilon_0 U^2 \frac{1}{\left(r - R_0 + \frac{S}{\tan(\varphi)} \right)^2 \cdot \varphi^2} \cdot \cos(\varphi) \quad (4.1.21)$$

Um brauchbare Aussagen zu gewinnen, müssen schließlich die Aufpunkte berechnet werden, also jene Punkte auf der r-Achse, von denen die Bögen ausgehen, welche Anfangs- und Endpunkt des Kegels schneiden. Aus einfachen geometrischen Überlegungen erhält man:

$$\begin{aligned} r_1 &= S \frac{1 - \cos(\varphi)}{\sin(\varphi)} + R_0 \\ r_2 &= r_1 + \frac{W}{\cos(\varphi)} \end{aligned} \quad (4.1.22)$$

Um das gewonnene Resultat zu überprüfen, wird das Randwertproblem numerisch gelöst. Die benutzten Parameter sind (Längen in willkürlichen Einheiten): $H=5-15$, $R=0.01-0.05$, $S=0.05-0.1$, Öffnungswinkel $\delta=10^\circ$. Aus den Lösungen wurde der Verlauf des Quadrates des elektrischen Feldes entlang der r-Achse extrahiert. Aus diesen Daten wurde durch Integration in Polarkoordinaten die wirkende Gesamtkraft erhalten. Der Wert dieses Integrals ist beispielhaft

für festes R_0 und variierendes S in Abbildung 4.8 dargestellt. Dabei entsprechen Symbole den numerisch gewonnenen Daten und Linien Anpassungen der in Gleichung (4.1.23) angegebenen Form.

$$I = A \ln \left(\frac{H}{S+B} \right) + C \quad (4.1.23)$$

Tatsächlich ist es die Variation der freien Parameter (A , B , C) sehr gering, was die Gültigkeit der durchgeführten Näherungen bestätigt.

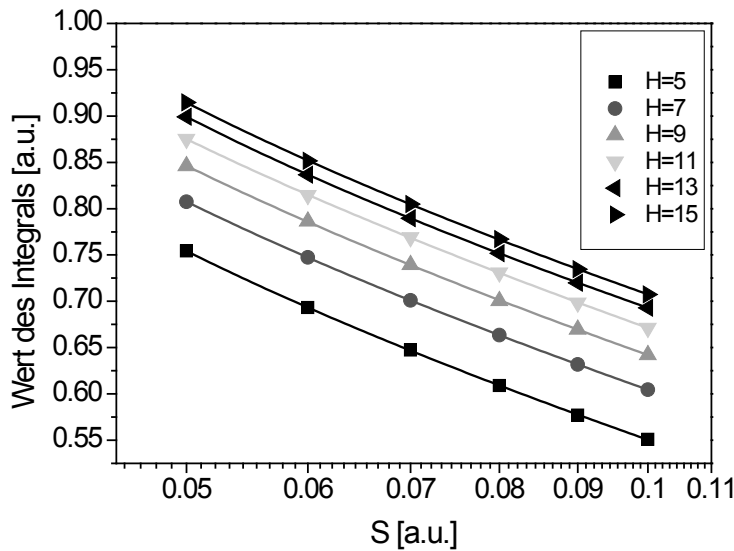


Abbildung 4.8: Gesamtkraft wirkend auf einen Kegelstumpf in Abhängigkeit von dessen Höhe und Abstand zur Oberfläche. Die numerisch gewonnenen Daten lassen sich sehr gut mit einem einfachen Modell beschreiben.

Die Transferfunktion ergibt sich nun durch Normierung auf die Gesamtkraft:

$$T(r) = \frac{F_z(r)}{\int_{r_1}^{r_2} 2\pi r F_z(r)} \quad (4.1.24)$$

Da das dabei auftretende Integral zwar analytisch lösbar ist, allerdings eine etwas unhandliche Form zeigt, soll hier die Transferfunktion in einer für übliche Parameter näherungsweise gültigen Gestalt angegeben werden. Dazu beachtet man, dass der Winkel ϕ im Bereich von 80° liegt, S in der Größenordnung von R_0 ist und die Höhe des Kegels H deutlich grösser als S ist. Man erhält, wenn der Winkel zwischen Oberfläche und Kegel durch den halben Öffnungswinkel δ des Kegels ersetzt wird:

$$T(r) = \frac{1}{2\pi} \frac{1}{\ln \left(\frac{H}{S+2R_0} \right)} \frac{1}{(r+S \tan(\delta) - R_0)^2} \quad (4.1.25)$$

Die Transferfunktion

4.1.3. Die Cantileverbasis

Offenbar ist es ausreichend, den effektiven Abstand der Oberfläche zur Cantileverbasis und somit das elektrische Feld als konstant anzunehmen. Es gilt somit:

$$\begin{aligned} L(r) &= \text{const} = H; \quad r > r_C \\ E(r) &= \frac{U}{H^2}; \quad r > r_C \end{aligned} \quad (4.1.26)$$

Zu klären bleibt allerdings, welchen Wert der kritische Radius r_C annimmt. Gemäß Gleichung (4.1.22) sollte gelten, dass r_C für kleine Apexradien und kleine Öffnungswinkel des Kegels etwa der Höhe H des Kegels entspricht. Das scheint widersprüchlich zu der intuitiven Annahme zu sein, dass der Einfluss der Basis für Radien grösser als die Weite des Kegels existent ist. Eine grobe Abschätzung des Wertes kann durch Anwendung des Dirichletschen Prinzips [28] gewonnen werden, was für das vorliegende Problem aussagt, dass die elektrostatische Energie minimal sein muss. Somit gilt folgende Bestimmungsgleichung:

$$\frac{\partial}{\partial r_C} \left(\int_0^{r_C} \frac{r \cdot (r+c)}{(r+c)^2} dr + \int_{r_C}^{W_{Basis}} \frac{r \cdot H}{H^2} dr \right) = 0 \quad (4.1.27)$$

In (4.1.27) ist c eine Konstante, die von der Separation S und dem Radius des Apex abhängt, und W_{Basis} die Breite der Cantileverbasis. Zum Aufstellen der Gleichung wurde angesetzt, dass die Energie über das Integral der Energiedichte, dem Quadrat des elektrischen Feldes, definiert ist. Die Durchführung der Ableitung liefert schließlich:

$$\begin{aligned} \frac{r_C}{(r_C+c)} &= \frac{r_C}{H} \\ H-c &= r_C \end{aligned} \quad (4.1.28)$$

Da die Konstante c in der Größenordnung von S und somit deutlich kleiner als H ist, erscheint die Annahme $r_C=H$ als gerechtfertigt.

4.1.4. Zusammenführung der Teilbereiche

Aus den vorausgegangenen Betrachtungen kann ein, stückweise definiertes, Modell für die effektiven Abstände L und somit für die Transferfunktion angegeben werden. Der Einfachheit halber sollen folgende, durchaus sinnvolle, Vereinfachungen angenommen werden:

1. Der Öffnungswinkel des Kegels ist klein. Somit beträgt der in den Betrachtungen genutzte Winkel ϕ etwa 90° .
2. Der Abstand S zwischen Spitze und Probe ist in derselben Größenordnung wie der Apexradius R_{Apex} und somit deutlich kleiner als die Höhe H des Kegels

Man kann also folgende Funktion für L angeben:

$$L(r) = \begin{cases} S + f_1(S, R_{Apex}) \cdot r^2; & 0 \leq r \leq S + 2R_{Apex} \\ r + f_2(S, R_{Apex}, \phi); & S + 2R_{Apex} \leq r \leq H \\ H; & H \leq r \leq W_{Basis} \end{cases} \quad (4.1.29)$$

Die Form der hier benutzten Funktionen f_1 und f_2 entnimmt man (4.1.15) und (4.1.19). Um die in (4.1.29) eingeführte Definition zu verifizieren, wurde das Randwertproblem numerisch gelöst und die Länge der elektrischen Feldlinien berechnet. Diese sollte deren Länge durch die Funktion L beschreibbar sein. Das Ergebnis dieser Prozedur für verschiedene Separationen S zeigt Abbildung 4.9.

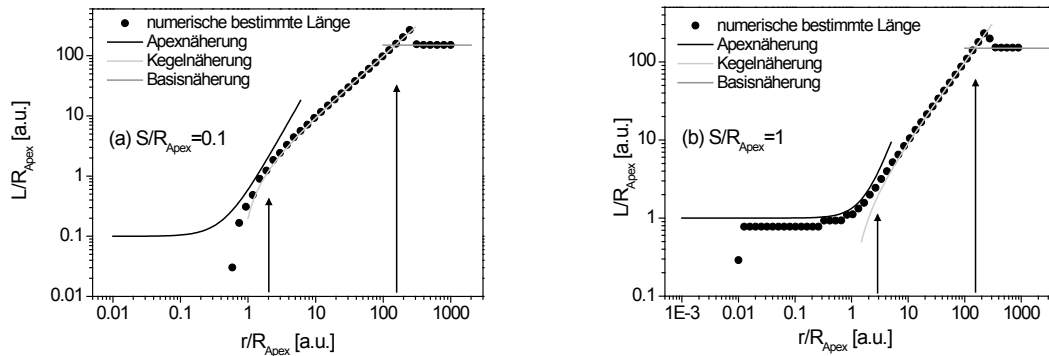


Abbildung 4.9: Numerisch bestimmte Länge der Feldlinien (Punkte) und Vergleich mit den Approximationen für kleine Separationen (a) und große (b). Sowohl die Modelle als auch die berechneten Grenzwerte zeigen gute Übereinstimmung mit den extrahierten Werten.

Offenbar lässt sich die numerisch bestimmte Länge der Feldlinien gut durch die analytische Formulierung annähern. Die auftretenden Abweichungen sind durch numerische Artefakte und zu grobe Abschätzung der Konstanten zu erklären. Sehr gut stimmen allerdings die durch Pfeile gekennzeichneten Intervallgrenzen überein. Werden die Parameter für die Funktion L über eine Kurvenanpassung bestimmt, ergibt sich eine fast vollständige Reproduktion der Daten, wie Abbildung 4.10 illustriert.

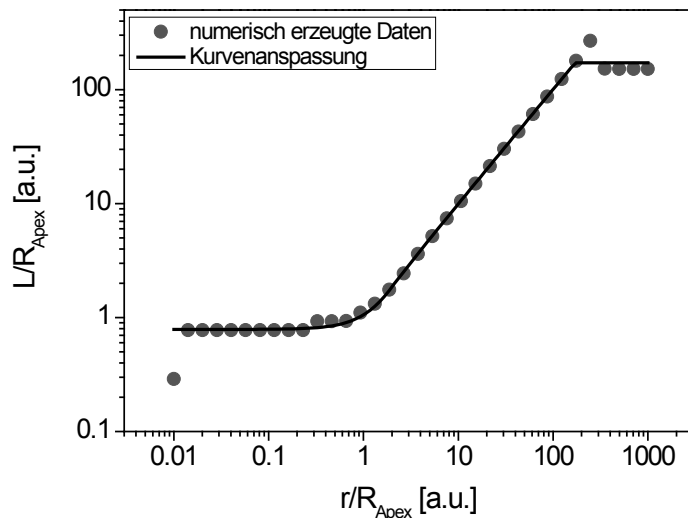


Abbildung 4.10: Kurvenanpassung an numerisch gewonnen Werte für die Länge der Feldlinien L . Bei Anpassung der freien Parameter ergibt sich beeindruckende Übereinstimmung.

Um von dieser, offenbar richtigen, Darstellung der Länge der Feldlinien auf die Transferfunktion schließen zu können, muss wiederum das Quadrat des elektrischen Feldes integriert werden:

Die Transferfunktion

$$T(r) = \frac{1}{L(r)^2} \cdot \frac{1}{\int_0^\infty \frac{2\pi r}{L(r)^2} dr} = \frac{1}{L(r)^2} \cdot \frac{1}{I} \quad (4.1.30)$$

Das auftretende Integral ergibt sich, unter Anwendung der üblichen Näherungen, zu:

$$I = 3\pi R_{Apex} \left(\frac{1}{S} - \frac{1}{S + 3R_{Apex}} \right) + 2\pi \ln \left(\frac{H}{S + 3R_{Apex}} \right) + \pi \left(\frac{W^2}{H^2} - 1 \right) \quad (4.1.31)$$

Die auftretenden Terme in (4.1.31) zeigen eine stark unterschiedliche Abhängigkeit von S : Während der durch den Apex bestimmte Wert etwa indirekt proportional zu S ist, zeigt der durch den Kegel bestimmte für moderate Werte von S einen festen Wert. Der Einfluss der Cantileverbasis ist für übliche Separationen und Verhältnisse von Weite zu Höhe vernachlässigbar, wie Abbildung 4.11 entnommen werden kann. Der Wert des Integrals, und somit die gesamte auf die Spitze wirkende Kraft, ist im dargestellten Bereich sehr gut durch Apex- und Kegelanteil beschreibbar.

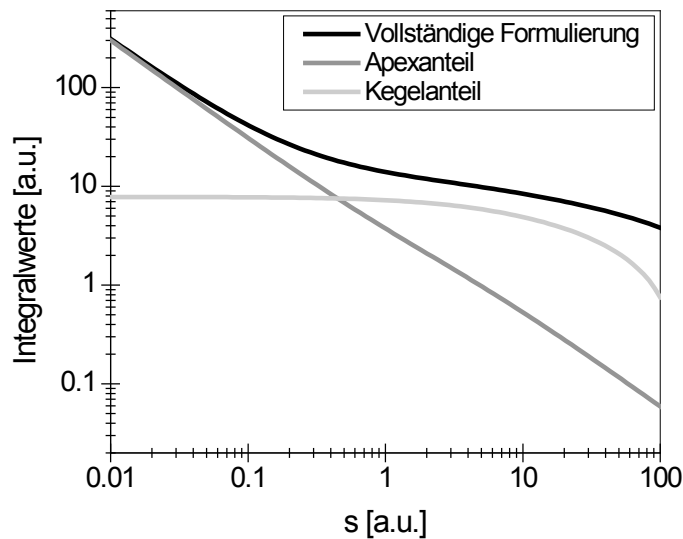


Abbildung 4.11: Vergleich der in (4.1.31) gefundenen Abhängigkeiten. Offenbar ist für die Gesamtkraft, welche dem Wert des Integrals bei vollständiger Formulierung entspricht, für kleine Separationen s lediglich der Apex beschreibend.

4.2 Numerische Bestimmung der Transferfunktion

Die systemspezifische Transferfunktion lässt sich numerisch auf verschiedenen Art und Weise bestimmen: Entweder durch numerische Nachstellung eines Experimentes oder durch direkte Bestimmung der Kraft-Ort-Abhängigkeit.

4.2.1 Numerische Nachstellung des Experimentes

Der Ablauf dieser Prozedur ist wie folgt:

- 1.) Nachdem die Geometrie des Systems modelliert ist, wird die Potentialverteilung auf der virtuellen Probenoberfläche durch eine geeignete Funktion beschrieben.
- 2.) Die auf die Spitze wirkende Kraft F wird in Abhängigkeit der an der Spitze angelegten Spannung V_T bestimmt.
- 3.) Es wird numerisch das Minimum V_{Min} der Funktion $F(V_T)$, hier durch quadratische Anpassung und analytische Auswertung der Koeffizienten, bestimmt.
- 4.) Abhängig von der angesetzten Potentialverteilung kann nun die Transferfunktion berechnet werden.

Es bieten sich zwei Funktionenklassen für die Potentialverteilung an: Stufenfunktionen und Deltafunktionen¹¹. Bei Verwendung von radialen Stufenfunktionen ergibt sich:

$$V_{Min}(r_0) = \int_0^{\infty} 2\pi r T(r) \cdot (1 - \Theta(r - r_0)) dr = \int_0^{r_0} 2\pi r T(r) dr \quad (4.2.1)$$

Offenbar gilt weiter:

$$\frac{dV_{Min}(r_0)}{dr_0} = 2\pi r_0 T(r_0) \Rightarrow T(r_0) = \frac{1}{2\pi r_0} \frac{dV_{Min}(r_0)}{dr_0} \quad (4.2.2)$$

Die Transferfunktion ist also durch numerische Auswertung der Ableitung des Minimums V_{Min} nach der Position des Sprungs r_0 bestimmbar.

Alternativ können, wie bereits auch erwähnt, Deltafunktionen genutzt werden, wobei sich ergibt:

$$V_{Min}(r_0) = \int_0^{\infty} 2\pi r T(r) \cdot \delta(r - r_0) dr = 2\pi r_0 T(r_0) \Rightarrow T(r_0) = \frac{V_{Min}(r_0)}{2\pi r_0} \quad (4.2.3)$$

Da numerisch tatsächlich nur Approximation der Deltadistribution zugänglich sind, sollen zwei übliche Näherungsfunktionen genauer untersucht werden. Als erstes erscheint es sinnvoll, eine „Box“-Funktion wie folgt zu definieren:

$$B(r, r_0, w) = \begin{cases} C; & r_0 - w < r < r_0 + w \\ 0; & \text{sonst} \end{cases} \quad (4.2.4)$$

Die Konstante C ergibt sich aus der Bedingung der Normierung des Integrals:

$$\int_0^{\infty} B(r, r_0, w) dr = 1 = \begin{cases} 2wC; & r_0 > w \\ (r_0 + w)C; & r_0 < w \end{cases} \Rightarrow \quad (4.2.5)$$

$$C = \frac{1}{w + \min(w, r_0)}$$

¹¹ Gemeint sind im Weiteren natürlich Repräsentationen der Deltadistribution: Funktionen, welche nur in einem kleinen Intervall von Null verschiedene, positive Werte liefern und integralnormiert sind.

Die Transferfunktion

Eine weitere Möglichkeit stellt die Nutzung von Gaußschen Glockenkurven dar, die wie folgt definiert werden:

$$G(r, r_0, w) = C \exp\left(-\frac{(r - r_0)^2}{w^2}\right) \quad (4.2.6)$$

Auch hier kann die Konstante C über die Normierungsbedingung bestimmt werden:

$$\begin{aligned} \int_0^\infty C \exp\left(-\frac{(r - r_0)^2}{w^2}\right) dr &= 1 = \frac{\sqrt{\pi} w C}{2} \left(\operatorname{erf}\left(\frac{r_0}{w}\right) + 1 \right) \Rightarrow \\ C &= \frac{2}{\sqrt{\pi} w} \frac{1}{\left(\operatorname{erf}\left(\frac{r_0}{w}\right) + 1 \right)} \end{aligned} \quad (4.2.7)$$

Um die Auswahl der geeigneten Variante zu motivieren, sollen die auftretenden Fehler abgeschätzt werden. Dazu werden die numerisch durchgeführten Schritte analytisch, unter der Annahme glatter Transferfunktionen, nachgestellt. Es ergeben sich folgende Zusammenhänge:

- a) Integrationsmethode: Gemäß Vorschrift werden zur Bestimmung der Transferfunktion am Punkt r_0 zwei Werte V_{Min} benötigt. Diese werden links- und rechtsseitig von r_0 berechnet, da die Verwendung des zentralen Differenzenquotienten eine höhere Genauigkeit liefert. Die Ermittlung dieses Differenzenquotienten ergibt schließlich:

$$T(r_0)_{app} = \frac{1}{2\pi r_0} \frac{\Delta V_{Min}(r_0)}{\Delta r_0} \approx \frac{1}{r_0} \frac{1}{2w} \int_{r_0-w}^{r_0+w} r T(r) dr = T(r_0)_{exakt} + c w^2 \quad (4.2.8)$$

Für die Herleitung von (4.2.8) wurde äquidistanter Abstand w zu r_0 der beiden Bestimmungspunkte angenommen und der Integrand um r_0 bis zur zweiten Ordnung entwickelt. Die auftretende Konstante c enthält dementsprechend Anteile der ersten und zweiten Ableitung von T .

- b) Bei Benutzung der Box-Funktionen ergibt sich offenbar derselbe funktionale Zusammenhang zwischen numerisch bestimmtem Wert und tatsächlichem, allerdings entspricht die Weite w hier der Weite der Box und nicht dem Abstand zwischen zwei Bestimmungspunkten.
- c) Die Nutzung der Glockenkurven erzeugt schließlich einen zusätzlichen Term, der wiederum Ableitungen von T enthält:

$$T(r_0)_{app} = T(r_0)_{exakt} + b w + c w^2 \quad (4.2.9)$$

Ausgehend von diesen Abschätzungen erscheint es also sinnvoll, die integrative Methode oder die Abtastung mittels Box-Funktion zur Darstellung der Transferfunktion zu wählen, da diese beiden äquivalente Ergebnisse liefern und die Abweichung, verglichen mit den Glockenkurven, geringer scheint. Um diese Vorhersage zu überprüfen, soll das Szenario numerisch nachgestellt

werden: Dazu wird die Transferfunktion gemäß (4.1.18) angenommen und die Integrale numerisch berechnet.

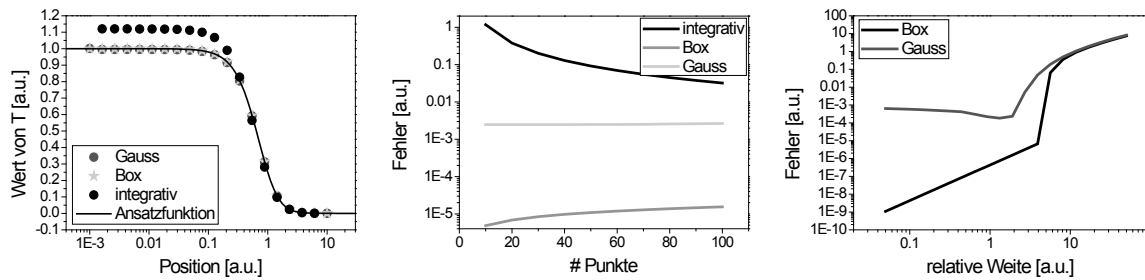


Abbildung 4.12: (Links) Vollständig numerischer Test zur Nutzbarkeit verschiedener Abläufe zur Bestimmung der Transferfunktion. Die integrative Methode erzeugt offenbar immer einen großen Fehler, die Methoden "Box" und "Gauß" arbeiten scheinbar zufriedenstellend (Mitte). Allerdings gibt es für eben diese eine deutliche Abhängigkeit der Güte von der Breite der Abtastfunktion (rechts).

Dieser, zugegeben vollständig theoretische, Ansatz liefert unter anderem das verblüffende Ergebnis, dass Abtastung mit „Box“-Funktionen am günstigsten ist. Aufgrund der allgemeinen Erfahrung bezüglich der Abbildung von Stufen jeglicher Art in numerischen Schemata erscheint dies nicht gerechtfertigt. Tatsächlich unterliegen starke Änderungen, sowohl in Koordinaten als auch in Werten, gewissen Dämpfungseinflüssen, welche nachfolgend besprochen werden sollen.

4.2.2 Abschätzung der intrinsischen Dämpfung des Laplace-Operators

Behandlung des Randwertproblems in zwei Dimensionen

Die zu behandelnden Randwertprobleme können für den Fall von fixierten Ladungen freier Gebiete wie folgt notiert werden:

$$\begin{aligned}\Delta V &= 0; (x, z) \in \Omega \\ V &= g(x, z); (x, z) \in \partial\Omega\end{aligned}\tag{4.2.10}$$

In (4.2.10) steht V für das elektrostatische Potential, x und z für die Koordinaten, Ω für das betrachtete Gebiet und die Funktion g gibt schließlich die Randwerte auf dem Rand $\partial\Omega$ an. Eine direkte analytische Lösung für beliebige g ist im Allgemeinen nicht einfach zugänglich, deshalb soll im Folgenden ein interessanter Prototyp definiert und gelöst werden.

Das Gebiet Ω soll ohne Einschränkung der Allgemeinheit folgende einfache Form annehmen:

$$\Omega = \{x \in [-B, B], z \in [0, H]\}\tag{4.2.11}$$

Auf den Rändern soll gelten:

$$\begin{aligned}V(x, z=0) &= \cos\left(\frac{2\pi x}{B}\right) \\ V(x, z=H) &= 0 \\ \frac{\partial V}{\partial x}(x=B, z) &= \frac{\partial V}{\partial x}(x=-B, z) = 0\end{aligned}\tag{4.2.12}$$

Die Transferfunktion

Offenbar genügt es, folgenden Ansatz für die gesuchte Lösung V anzunehmen:

$$V(x, z) = A(z) \cos\left(\frac{2\pi x}{B}\right) \quad (4.2.13)$$

Die Randbedingungen am unteren und den äußeren Rändern werden automatisch erfüllt, zur Bestimmung der Amplitudenfunktion A wird der Ansatz in (4.2.10) eingesetzt:

$$\begin{aligned} \Delta V &= \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right) V = -A(z) \cos\left(\frac{2\pi x}{B}\right) \frac{4\pi^2}{B^2} + A''(z) \cos\left(\frac{2\pi x}{B}\right) = 0 \\ \Leftrightarrow A''(z) - A(z) k^2 &= 0 \end{aligned} \quad (4.2.14)$$

Die Lösung von (4.2.14) ist einfach und ergibt sich, nach Anpassung der Konstanten an die Randbedingungen zu:

$$A(z) = \frac{\exp(-kz)}{1 - \exp(-2kH)} + \frac{\exp(kz)}{1 - \exp(2kH)} \quad (4.2.15)$$

Es lassen sich zwei asymptotische Fälle betrachten: Für große H entfällt der zweite Term, was eine rein exponentielle Abnahme der Amplitudenfunktion verursacht während für kleine H die Funktion fast linear in z erscheint, wie Abbildung 4.13 verdeutlicht:

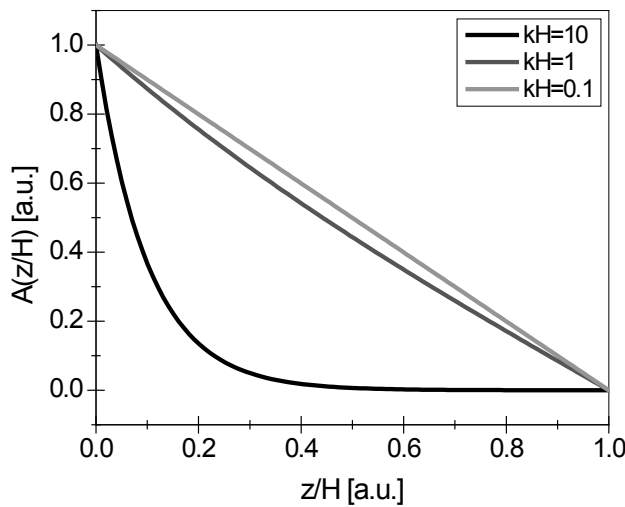


Abbildung 4.13: Verlauf der Amplitudenfunktion A für verschiedenen Höhen H des betrachteten Gebietes. Während für große H ein exponentieller Zusammenhang zwischen Koordinate und Amplitude besteht (schwarze Spur), ist der Zusammenhang für kleine Höhen linear (hellgraue Spur).

Einsichtiger wird das analytische Verhalten bei Normierung der Variablen:

$$\begin{aligned} \tilde{z} &= \frac{z}{H}, \quad \tilde{k} = kH \\ \Leftrightarrow A(\tilde{z}) &= \frac{\exp(-\tilde{k}\tilde{z})}{1 - \exp(-2\tilde{k})} + \frac{\exp(\tilde{k}\tilde{z})}{1 - \exp(2\tilde{k})} \end{aligned} \quad (4.2.16)$$

Offenbar ist der kritische Parameter in (4.2.16) das Produkt aus normierter z-Koordinate und normierter Wellenzahl oder, anschaulicher, das Verhältnis von z-Koordinate und Wellenlänge der Kosinusfunktion auf dem unteren Rand. Wie bereits erwähnt, nimmt die Amplitudenfunktion für große Ausdehnungen des Gebietes, was durchaus praktische Bedeutung hat, eine einfachere Form an¹²:

$$A(\tilde{z}) = \exp(-\tilde{k}\tilde{z}) \quad (4.2.17)$$

Aus (4.2.17) lassen sich zwei wichtige Schlussfolgerungen ziehen:

- 1.) Alle Spektralkomponenten einer Potentialverteilung auf dem unteren Rand werden exponentiell mit dem Abstand zu diesem Rand gedämpft.
- 2.) Die Abklingkonstanten entsprechen den Wellenzahlen. Dementsprechend werden kurze Wellenlängen schneller gedämpft als längere.

Interessanterweise kann das Ergebnis aus (4.2.17) auch wie folgt erhalten werden:

$$\int_{-\infty}^{\infty} \cos(ka) T(x, z, a) da = V(x, z) = A(z, k) \cos(kx) \quad (4.2.18)$$

Motiviert wurde diese Darstellung durch die Möglichkeit, Randwertprobleme für die Laplacegleichung durch Randintegrale mit geeigneten Greenschen Funktionen, hier T , zu erhalten [12]. Offenbar stellt die Amplitudenfunktion $A(z, k)$ den Fourierkoeffizienten von T bezüglich k dar, wodurch sich T bestimmen lässt zu:

$$T(x, z, a) = \frac{1}{\pi} \frac{z}{(x-a)^2 + z^2} \quad (4.2.19)$$

Nachdem offensichtliche Grenzübergänge, das heißt die Erweiterung des in (4.2.11) definierten Gebietes auf den Halbraum, vorgenommen wurden, lässt sich also verallgemeinert festhalten:

Das Potential $V(x, z)$ lässt sich durch Faltung der Potentialverteilung $V(x, 0)$ auf dem unteren Rand mit einer Lorentzfunktion der Breite z erhalten. Dieses Ergebnis gilt für alle Randfunktionen, da diese nach ihren Spektralkomponenten zerlegt werden können.

Behandlung des Randwertproblems in drei Dimensionen

Eine Erweiterung der Theorie auf drei Dimensionen stellt sich etwas schwieriger dar: Während die einfachen Lösungsansätze (4.2.13)-(4.2.14) quasi übernommen werden können, scheint die Erschließung einer geschlossenen analytischen Formulierung zu komplex. Analog zu den bereits ausgeführten Überlegungen formuliert man:

¹² Es ist möglich, Faltungsformeln für nicht-unendliche Höhe des Gebietes zu bestimmen. Eine kurze Erläuterung findet sich am Ende dieses Paragraphen.

Die Transferfunktion

$$\Delta V = -A(z) \cos\left(\frac{2\pi x}{B_x}\right) \cos\left(\frac{2\pi y}{B_y}\right) \cdot \left(\frac{4\pi^2}{B_x^2} + \frac{4\pi^2}{B_y^2}\right) + A''(z) \cos\left(\frac{2\pi x}{B_x}\right) \cos\left(\frac{2\pi y}{B_y}\right) = 0$$

$$\Leftrightarrow A''(z) - A(z)k_{xy}^2 = 0$$

(4.2.20)

In (4.2.20) wurde das Modell auf den Quader der Kantenlängen $2B_x, 2B_y, H$ erweitert und die Randwertfunktion als Produkt zweier eindimensionaler Funktionen definiert. Die hier erscheinende effektive Wellenzahl k_{xy} ergibt sich intuitiv gemäß:

$$k_{xy}^2 = \frac{4\pi^2}{B_x^2} + \frac{4\pi^2}{B_y^2} = k_x^2 + k_y^2 \quad (4.2.21)$$

Somit ist der Verlauf der Amplitudenfunktion entsprechend (4.2.15), wenn k durch k_{xy} ersetzt wird.

Die Darstellung der Transferfunktion $T_{3D}(x, y, z, a, b)$ ist nicht trivial, sehr wahrscheinlich nimmt sie aber folgende Gestalt an:

$$T_{3D}(x, y, z, a, b) = \frac{1}{2\pi} \frac{z}{\left((x-a)^2 + (y-b)^2 + z^2\right)^{\frac{3}{2}}} \quad (4.2.22)$$

Plausibel wird (4.2.22) durch folgende Überlegungen:

- 1.) Eine Integration von $T_{3D}(x, y, z, a, b)$ über b , also für von y unabhängige Randwertfunktionen, ergibt (4.2.19). Diese Bedingung muss erfüllt sein, da für diesen speziellen Fall die zweidimensionalen Formulierungen greifen müssen.
- 2.) $T_{3D}(x, y, z, a, b)$ ist ebenfalls als Richtungsableitung der Fundamentallösung der Poissongleichung darstellbar.

Verallgemeinert kann somit angegeben werden, wenn A den unteren Rand und f die Randwertfunktion benennen:

$$V(x, y, z) = \int_A T(x, y, z, a, b) f(x+a, y+b) dA \quad (4.2.23)$$

Zusammenfassend bleibt also zu sagen:

- a) Sowohl in zwei als auch in drei Dimensionen werden die Spektralkomponenten der Randwertfunktionen exponentiell gedämpft, wobei die Dämpfungsparameter ihren (effektiven) Wellenzahlen entsprechen.
- b) Es ist möglich, analytische Transferfunktionen T anzugeben, die eine integrale Bestimmung des Potentials $V(x, y, z)$ gemäß (4.2.23) erlauben.

Erweiterung der Theorie auf zweidimensionale Streifen

Ist die Höhe der betrachteten Domäne nicht unendlich, beziehungsweise die betrachteten z -Koordinaten liegen nahe am oberen Rand, gilt die Näherung (4.2.17) nicht mehr und die Faltung (4.2.18) kann nicht benutzt werden.

Mit Hilfe konformer Abbildung kann allerdings ein Streifen in die obere Halbebene abgebildet werden. Die dafür gefundenen Integralgleichungen werden dann zurück transformiert, was folgendes Ergebnis liefert (die Koordinaten sind auf die Höhe H des Streifens normiert,[29]):

$$V(x, z) = \frac{\sin(\pi z)}{2} \int_{-\infty}^{\infty} \frac{g(x-t)}{\cosh(\pi t) - \cos(\pi z)} dt \quad (4.2.24)$$

In (4.2.24) gibt $g(x)$ wiederum die Randwerte auf dem unteren Rand $z=0$ an. Da das auftretende Integral im Allgemeinen nicht analytisch lösbar ist, sollen zwei Zugänge zur Plausibilitätsanalyse gewählt werden:

- 1.) Eine Entwicklung für kleine z (oder sehr große H): Eine Taylorentwicklung von (4.2.24) um $z=0$ liefert:

$$V(x, z) \approx \frac{1}{2} \int_{-\infty}^{\infty} \frac{\pi z \cdot g(x-t)}{\cosh(\pi t) - 1} dt \quad (4.2.25)$$

Für ebenfalls kleine t stimmt diese Formulierung mit der in (4.2.18) überein, wie Abbildung 4.14 verdeutlicht. Offenbar ist der Verlauf für weite Bereiche identisch, die Abweichungen für größere Werte von t dürften aufgrund der kleinen Funktionswerte kaum ins Gewicht fallen.

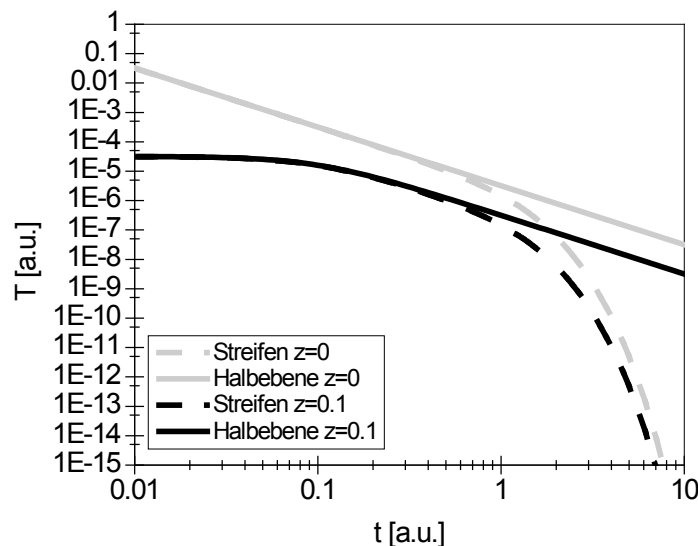


Abbildung 4.14: Vergleich der Transferfunktionen für streifenförmige Gebiete (schwarze Spuren) und die positive Halbebene (graue Spuren). Die für $z=0$ angegebenen Funktionen entsprechen dabei Näherungen für kleine Werte der Koordinaten (t, z).

Die Transferfunktion

- 2.) Eine numerische Auswertung des Integrals für das Beispielproblem (4.2.12) liefert die in Abbildung 4.15 dargestellten Verläufe. Tatsächlich ist der Grafik zu entnehmen, dass die durch (4.2.24) berechneten Werte (Punkte) denen aus (4.2.16) (Linien) entsprechen. Somit erscheint die Gültigkeit der Integralformel für streifenförmige Gebiete im Rahmen dieser Arbeit als gegeben.

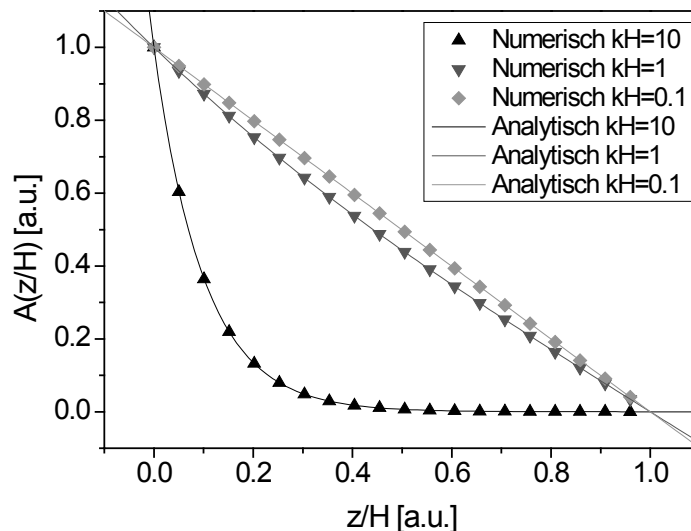


Abbildung 4.15: Vergleich der Ergebnisse der Integralformel für streifenförmige Gebiete (Punkte) mit den analytischen Ergebnissen für das Modellproblem (Linien). Die Übereinstimmung ist sehr gut, eine Kurvenanpassung der Form (4.2.16) an die Datenpunkte liefert hervorragende Deckung des freien Parameters kH .

4.2.3 Abschätzung der numerischen Dämpfung bei der Behandlung elektrostatischer Randwertprobleme

Es ist hinlänglich bekannt, dass die numerische Behandlung jeglicher Problemstellungen Fehler birgt, welche durch die notwendige Diskretisierung entstehen. Zum einen werden Abtastartefakte erzeugt, da vorgegebene Funktionen nur an definierten Punkten erfasst werden und zum anderen werden die notwendigen Differentialoperatoren durch diskrete Äquivalente ersetzt, die meist durch Approximation ermittelt werden. Exemplarisch sollen die auftretenden Probleme am Beispiel einer zweidimensionalen Laplacegleichung mit gemischten Randwerten erklärt werden. Die Problemstellung und Geometrie sind in Abbildung 4.16 skizziert.

Das Vorgehen ist wie folgt: Das Gebiet wird in äquidistante Gitterpunkte in x, y zerlegt und die so erhaltenen Punkte zu Vektorform umsortiert. Anschließend wird das diskrete Äquivalent des Laplaceoperators erzeugt. Die Randwerte werden in die, eigentlich nur „Null“-Einträge enthaltende, rechte Seite eingefügt. Die Lösung des so aufgestellten Gleichungssystems ist relativ einfach zu handhaben und wird über erhältliche Software durchgeführt.

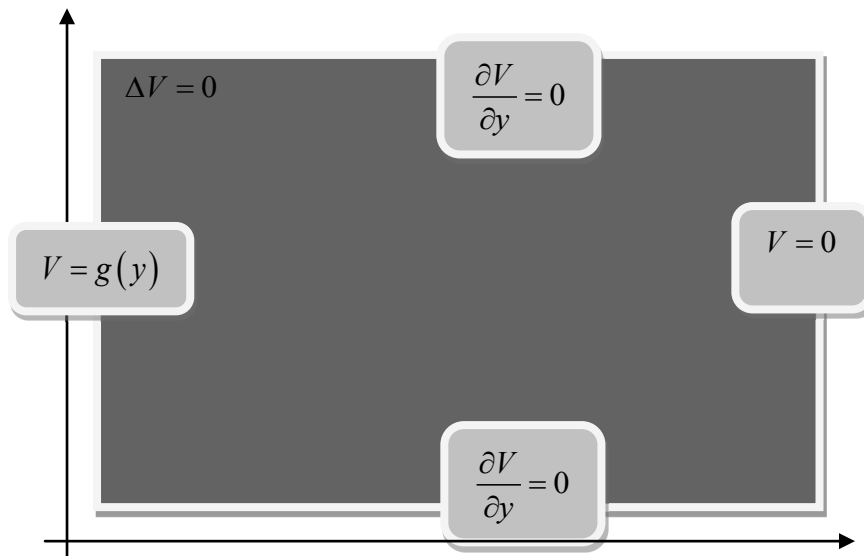


Abbildung 4.16: Struktur der Domäne zur numerischen Lösung eines Standardproblems

Mathematisch formuliert ergibt sich:

$$\mathbf{L}\vec{V} = \vec{r} \quad (4.2.26)$$

$$\vec{r} = (g(y_1), g(y_2), \dots, g(y_N), 0, \dots, 0)$$

Um brauchbare und vor allem vergleichbare Aussagen zu erhalten, werden Cosinusfunktionen für die Randwerte angesetzt:

$$g(y) = \cos\left(\frac{2\pi}{\lambda} y\right); \quad \lambda = \frac{2(N_y - 1)}{m} \quad (4.2.27)$$

Die Wellenlänge λ wird dabei in Abhängigkeit der Anzahl der Punkte in y-Richtung N_y so gewählt, dass die Symmetriebedingung an den entsprechenden Rändern erfüllt wird. Der Parameter m gibt also an, wie oft die halbe Wellenlänge in das Gebiet passt. Nach Umsortieren der vektorförmigen Lösung ergeben sich beispielsweise die in Abbildung 4.17 gezeigten Bilder.

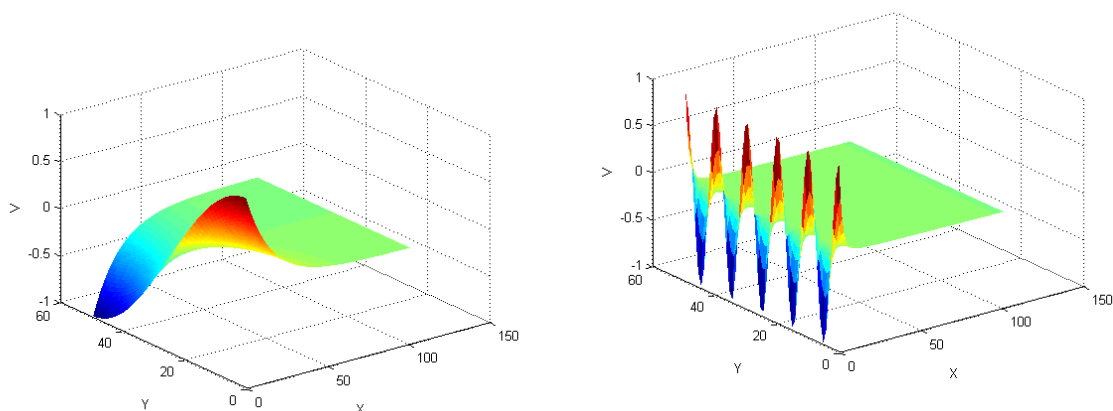


Abbildung 4.17: Numerische Lösung des in (4.2.26) und (4.2.27) definierten Randwertproblems mit Hilfe der Methode der finiten Differenzen. Sowohl für kleine Wellenlängen ($m=1$, links) als auch für große Wellenlängen ($m=10$, rechts) erscheint die erwartete Struktur.

Die Transferfunktion

Die numerisch generierten Lösungen spiegeln die analytisch gewonnenen Erwartungen (Paragraph 4.2.2) wieder, zumindest bei visueller Inspektion. Genauere Aussagen erhält man, wenn der Verlauf der Lösung an die analytische Repräsentation angepasst wird und die entsprechenden Parameter extrahiert werden.

Dazu werden die erzeugten Lösungswerte parallel zur x-Achse, also für feste y , an Exponentialfunktionen angepasst, wobei die signifikanten Parameter offenbar Amplitude A und Dämpfungsparameter β sind:

$$V(x, y_0) = A(y_0) \exp(-\beta x) \quad (4.2.28)$$

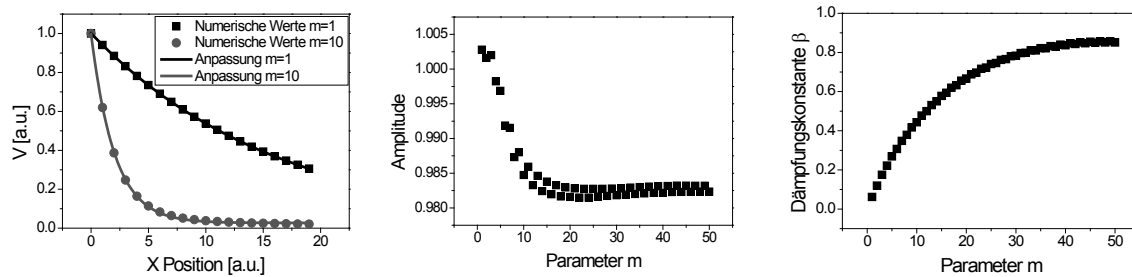


Abbildung 4.18: Kurvenanpassungen der numerisch gewonnenen Lösung an die analytische Form. Während der Verlauf der Funktionen zufriedenstellend erscheint (links) zeigen Amplitudengang (Mitte) und extrahierter Dämpfungsparameter (rechts) eine deutliche Abhängigkeit vom benutzten Skalierungsparameter m , welcher indirekt proportional zu Wellenlänge ist.

Für feste Anzahl der Punkte in den Dimensionen (hier willkürlich 50 und 100) ergeben sich die in Abbildung 4.18 dargestellten Zusammenhänge: Der Verlauf der Kurven entspricht im Rahmen der numerischen Genauigkeit den Erwartungen, allerdings weichen Amplitudengang und Wert des Dämpfungsparameters doch merklich von den Vorhersagen ab. Entgegen der Erwartung ist die extrahierte Amplitude nicht konstant eins, sondern sinkt nichtlinear mit der Wellenlänge, hier charakterisiert durch den Parameter m . Dieses Verhalten lässt sich durch eine verallgemeinerte Form des Abtasttheorems [30] erklären: Wenn die Wellenlänge einen bestimmten Wert unterschreitet, das heißt die Anzahl der Abtastpunkte pro Wellenlänge zu gering wird, kann die Randwertfunktion nicht mehr mit genügender Genauigkeit abgebildet werden. Ausgehend von Abbildung 4.18 (Mitte) kann man, durchaus empirisch und ohne analytische Herleitung, einen kritischen Parameter $m_k=10$ definieren. Für Parameter grösser als dieser Wert bleibt die extrahierte Amplitude etwa 2% unter der erwarteten. In absoluten Zahlen ergibt sich somit die Bedingung, dass etwa 10 Abtastpunkte pro Wellenlänge für zufriedenstellende Genauigkeit nötig sind.

Der funktionale Zusammenhang zwischen Dämpfungsparameter und Wellenlänge ist leider nicht intuitiv zu erklären. Ein einfaches Modell kann allerdings zumindest den Verlauf qualitativ erläutern. Geht man davon aus, dass genügend Stützstellen parallel zum linken Rand vorhanden sind, um die Cosinusfunktion abzubilden, erscheint folgender Ansatz für einen beliebigen Punkt im Inneren des Gebietes sinnvoll:

$$V(x, y) = A \exp(-\beta x) \cos(ky) \quad (4.2.29)$$

Ausformulieren von (4.2.26) und Einsetzen des Ansatzes unter Einführung der Gitterweite Δ liefert:

$$\begin{aligned} V(x - \Delta, y) - 4V(x, y) + V(x + \Delta, y) + V(x, y - \Delta) + V(x, y + \Delta) = 0 \\ A \exp(-\beta(x - \Delta)) \cos(ky) - 4A \exp(-\beta x) \cos(ky) + A \exp(-\beta(x + \Delta)) \cos(ky) + \dots \\ + A \exp(-\beta x) \cos(k(y - \Delta)) + A \exp(-\beta x) \cos(k(y + \Delta)) = 0 \end{aligned} \quad (4.2.30)$$

Eine Anwendung von Additionstheoremen ermöglicht folgende Darstellung:

$$A \exp(-\beta x) \cos(ky) [2 \cosh(\beta \Delta) - 4 + 2 \cos(k \Delta)] = 0 \quad (4.2.31)$$

Offenbar muss zur Erfüllung der (Gitter)-Laplacegleichung der Term in den eckigen Klammern konstant null sein, was wiederum erlaubt, einen funktionalen Zusammenhang zwischen Wellenlänge und Dämpfung anzugeben:

$$\beta = \frac{1}{\Delta} a \cosh(2 - \cos(k \Delta)) = \frac{1}{\Delta} a \cosh\left(2 - \cos\left(\frac{2\pi}{\lambda} \Delta\right)\right) \quad (4.2.32)$$

Erfreulicherweise ergibt eine Taylorentwicklung von (4.2.32) für kleine Δ das analytische erwartete Ergebnis $\beta=k$.

Eine gemeinsame Darstellung der extrahierten Dämpfungskonstanten und einer (skalierten) Funktion gemäß (4.2.32) (Abbildung 4.19) zeigt leider keine zufriedenstellende Deckung, vielmehr erscheint es besser, stückweise lineare Abhängigkeit anzunehmen.

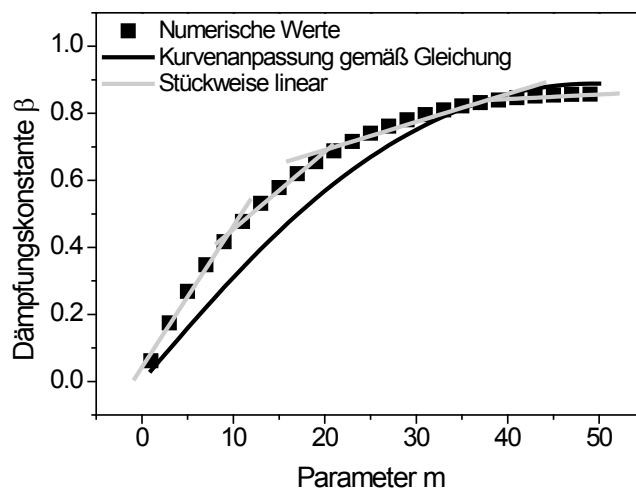


Abbildung 4.19: Modellierung der numerisch bestimmten Dämpfungskonstanten bei Benutzung der Methode der finiten Differenzen. Ein analytisch motiviertes Modell (schwarze Kurve) kann den Verlauf der Daten relativ gut widerspiegeln, allerdings erscheint die, relative simple, Annahme von Abtastartefakten in Form von Wellenlängenvervielfachung besser geeignet (graue Kurven).

Die numerisch bestimmten Anstiege verhalten sich dabei gemäß 1:0.5:0.2:0.03, weisen auf eine mögliche Vervielfachung der Grundwellenlänge hin. Im Rahmen dieses Ansatzes wird also,

Die Transferfunktion

abhängig vom Parameter m und somit von der Anzahl der Abtastpunkte pro Wellenlänge, nicht die ursprüngliche Funktion abgebildet, sondern Funktionen mit größerer Wellenlänge. Diese haben dann geringere Dämpfungskonstanten, was die Ergebnisse reflektieren. Der kritische Parameter m_k liegt auch hier bei 10, was in etwa 10 Punkten pro Wellenlänge entspricht.

Zusammenfassend kann formuliert werden, dass für genügend große Anzahl von Stützstellen die Dämpfung im diskretisierten Modell identisch zu der analytisch hergeleiteten ist. Für eine zu geringe Anzahl wird sie allerdings zu schwach abgeschätzt, was unweigerlich zu falschen Lösungen führt. Der generelle Trend der stärkeren Dämpfung kleiner Wellenlängen bleibt aber erhalten.

Im Weiteren ist davon auszugehen, dass die in den Simulationen modellierte Spitze möglicherweise nicht den Einfluss der definierten „Box“-Potentiale spürt. Vielmehr ist das tatsächlich abgebildete Potential eine geglättete Version des Ansatzes, was natürlich das analytisch motivierte Auswerten quasi ad absurdum führt. Wie vermutet, sollte, sowohl aufgrund numerischer als auch intrinsischer Dämpfung, eine glatte Abtastfunktion bevorzugt werden, etwa in Form der „Gauß“-Funktion. Diese benötigt nämlich weniger große Spektralanteile, wodurch ihre Dämpfung geringer ausfällt.

4.2.4 Beispiele für numerisch bestimmte Transferfunktionen

Ein typisches Beispiel für eine numerisch mit Hilfe der Abtastmethode mit Gaußfunktion bestimmte Transferfunktion zeigt Abbildung 4.20.

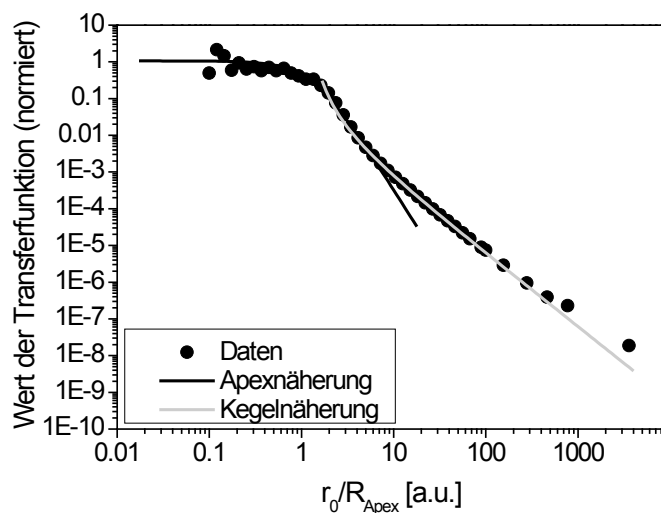


Abbildung 4.20: Numerisch bestimmte Transferfunktion für eine Spitze mit typischen Parametern (Symbole). Offenbar lässt sich die Funktion über weite Bereiche gut mit den theoretischen Überlegungen vereinbaren (Linien). Der Einfluss der Cantileverbasis auf die Messwerte erscheint gering, da die Näherungsfunktionen für Apex und Kegel bereits zufriedenstellende Übereinstimmung erzeugen.

Augenscheinlich lassen sich drei Bereiche im Verlauf unterscheiden:

1. Im Bereich $r < 2R_{Apex}$ ist die Funktion in doppeltlogarithmischer Darstellung schwach von r abhängig. Das entspricht dem Verlauf einer Lorentzfunktion wie in Gleichung (4.1.18) definiert.
2. Im Bereich $2R_{Apex} < r < 100R_{Apex}$ ist der Verlauf etwa indirekt proportional zum Quadrat des Radius r . Das entspricht der in Gleichung (4.1.25) angegebenen Funktion, welche den kegelförmigen Teil beschreibt.
3. Für sehr große Radien wird der Verlauf zunehmend flacher und kann, immer die numerischen Fehler bedenkend, als konstant angenommen werden. Das entspricht dem Einfluss der Basis.

Da der Einfluss der Cantileverbasis sehr gering ausfällt, soll im Weiteren das Abtasten der Transferfunktion nur für den Bereich erfolgen, in dem Apex und kegelförmiger Teil die Wechselwirkung dominieren. Das hat zum Einen numerische Vorteile und zum Anderen ist ohnehin zu erwarten, dass die Auflösung von realen Experimenten nur unwesentlich durch weiter vom Apex entfernte Bereiche bestimmt wird.

Abbildung 4.21 zeigt numerisch bestimmte Transferfunktionen für festen Radius des Apex (hier 10 nm) und verschiedene Separationen S . Deutlich erkennbar sind zwei wichtige Einflüsse: Zum einen sinkt der maximale Wert merklich mit der Separation und der Übergang zwischen den verschiedenen Bereichen verschiebt sich zu höheren Radien. Deutlicher wird diese Verschiebung bei Auftragung in auf das Maximum normierter Darstellung (rechts). Um greifbare, und vor allem mit den Modellen vergleichbare, Aussagen zu gewinnen, können sowohl das Maximum als auch ein gewisser Schwellradius r_c extrahiert werden. Der Schwellradius kann dabei entweder als der Punkt definiert werden, an dem der Wert der Transferfunktion einen bestimmten Bruchteil des Wertes für $r=0$ unterschreitet oder aber als der Wert, bei dem der zu erwartende Messwert sich bei Ausdehnungsänderung einer Inhomogenität nicht mehr ändert, das Kraftverhältnis f also einen bestimmten Wert erreicht.

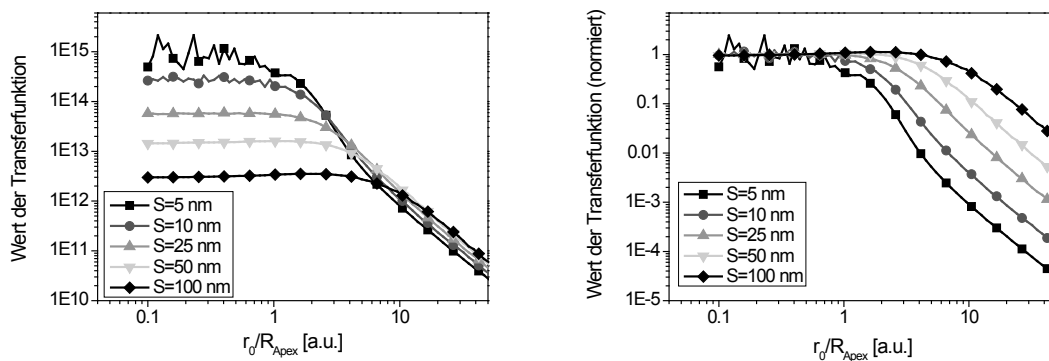


Abbildung 4.21: Numerisch bestimmte Transferfunktionen für verschiedene Separationen und festen Radius des Apex (links). Die normierte Darstellung (rechts) lässt kritische Parameter und Formänderungen deutlicher erkennen.

Formelmäßig lassen sich diese Schwellen, $r_{c,T}$ und $r_{c,F}$ wie folgt definieren:

$$\frac{T(r_{c,T})}{T(0)} = \frac{1}{16}$$

$$\frac{\int_0^{r_{c,F}} 2\pi r T(r) dr}{\int_0^\infty 2\pi r T(r) dr} = f \quad (4.2.33)$$

Die Konstante „16“ wurde dabei an Anlehnung an die Modelle gewählt (siehe Seite 87), prinzipiell kann auch jeder andere Wert genutzt werden. Die in (4.2.33) auftretenden Integrale müssen natürlich numerisch ausgewertet werden, die obere Grenze ist damit der maximal betrachtete Radius.

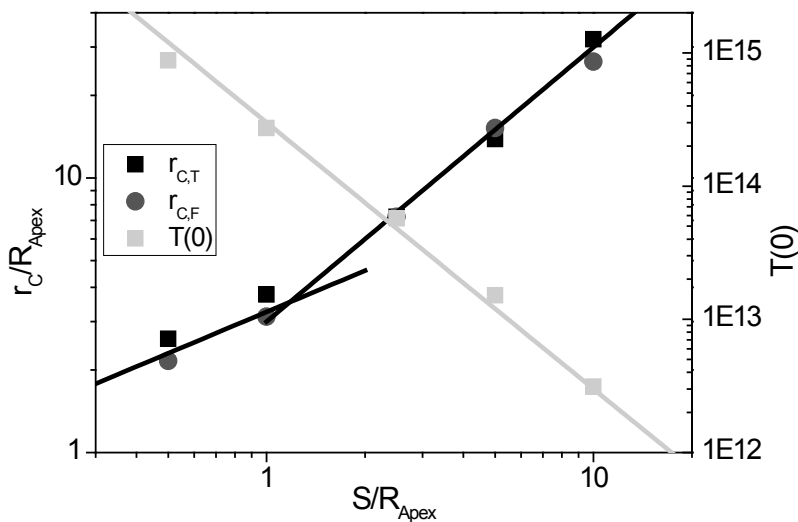


Abbildung 4.22: Darstellung von aus den Werten in Abbildung 4.21 bestimmten kritischen Radien und des maximalen Wertes der Transferfunktion. Offensichtlich lassen sich die Daten gut mit den theoretischen Vorhersagen des Apexmodells beschreiben, was im Endeffekt darauf hinweist, dass der Apex das limitierende Element ist.

Die in Abbildung 4.22 illustrierten Abhängigkeit (der Wert für das Kraftverhältnis wurde willkürlich auf $f=50\%$ festgelegt) der beschreibenden Parameter (Symbole) decken sich gut mit den Vorhersagen des Apexmodells: Für Separationen kleiner als der Apexradius steigt der Schwellradius proportional zur Wurzel der Separation, für größere Werte direkt proportional. Auch der maximale Wert fällt, wie dem Modell entnehmbar, etwa mit dem Quadrat der Separation ab. Somit kann vorsichtig formuliert werden, dass die Auswertung der numerisch bestimmten Transferfunktion ebenfalls eine Limitierung der Auflösung des Systems durch den Apex ergibt. Um diese Aussage zu untermauern, sollen parametrische Studien unternommen werden. Der wichtigste Parameter dabei ist offensichtlich der Apexradius. Weiter ist anzunehmen, dass der Öffnungswinkel des Kegels eine Rolle spielt. Dieser definiert nämlich den Punkt des Übergangs zwischen Apex und Kegel. Die Höhe des Kegels sollte eine untergeordnete Rolle spielen und wird an dieser Stelle nicht behandelt.

Ein Vergleich einer sehr scharfen mit einer sehr stumpfen Spitze (Abbildung 4.23, links) ergibt, dass es keine prinzipiellen Unterschiede ergibt. Die Kurvenanpassungen greifen bei beiden Spuren gut, die Weiten der Anpassungsfunktionen entsprechen den Erwartungen, sie sollten etwa bei dem Verhältnis von Separation zu Apexradius liegen.

Eine Darstellung der extrahierten kritischen Radien für unterschiedliche Geometrien, bestimmt direkt über die Transferfunktionen, zeigt Abbildung 4.23 (rechts). Obwohl eine leichte Variation der Daten zu entnehmen ist, liegen doch alle Werte sehr nahe an den einfachen Modellfunktionen. Die Einflüsse von Apexradius und Öffnungswinkel sind demnach im untersuchten Parameterbereich vernachlässigbar, es lassen sich allgemeine Vorschriften für das Auflösungsvermögen finden.

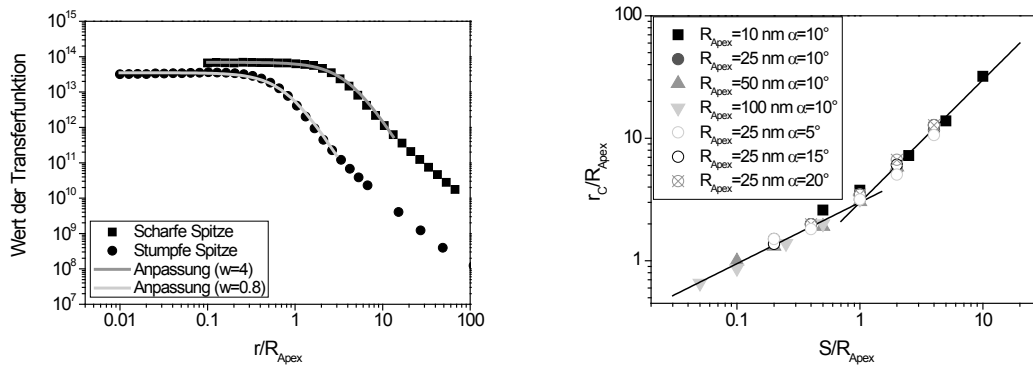


Abbildung 4.23: (Links) Numerisch bestimmte Transferfunktionen für eine scharfe Spitze ($R_{Apex}=10\text{nm}$, $\alpha=5^\circ$) und eine stumpfe ($R_{Apex}=100\text{nm}$, $\alpha=20^\circ$). Die Kurvenanpassungen ergeben für beide Kurven gute Übereinstimmung, die Verläufe sind sehr ähnlich. Extrahierte kritische Radien für Spitzen mit verschiedenen Apexradien und Öffnungswinkeln. Alle Werte liegen gehäuft um eine universelle Kurve, was die Angabe einer allgemein gültigen Auflösungsangabe motiviert.

Diese universelle Aussage, bestätigt sowohl durch numerische Simulation als auch einfache analytische Modellierung, lautet: In üblichen Bereichen von geometrischen Parametern und Separationen wird die Auflösung eines EFM-Experiments maßgeblich durch das Verhältnis von Separation zu Apexradius bestimmt. Die radiale Transferfunktion nimmt die Form einer abgewandelten Lorentzfunktion an, Weite und Amplitude dieser Funktion skalieren wiederum nur mit dem Verhältnis s zwischen Separation und Apexradius. Es gelten folgende, aus den numerischen Daten empirisch bestimmten, Zusammenhänge:

$$w(s) = \begin{cases} 3\sqrt{s}; & s < 1 \\ 3s; & s > 1 \end{cases} \quad (4.2.34)$$

4.2.5 Bestimmung der Transferfunktion durch Abbildung der Kraft

Wie eingangs in diesem Kapitel erwähnt, sollte angenommen werden können, dass eine direkte Abbildung des Kraftverlaufs auf der unteren Begrenzung der numerischen Domäne Zugang zur Transferfunktion bietet. Tatsächlich leidet dieser, auf Symmetrieüberlegungen beruhende Ansatz, unter der Tatsache, dass die Feldlinien gekrümmt sind. Demnach sind die vertikalen Komponenten des elektrischen Feldes an der Oberfläche der modellierten Spitze und an dem

Die Transferfunktion

korrespondierenden Punkt der unteren Begrenzung nicht identisch, was im Endeffekt eine leichte Änderung der vertikalen Kraftkomponente verursacht.

Eine vergleichende Darstellung der auf beiden beschriebenen Wegen erhaltenen Transferfunktionen zeigt Abbildung 4.24. Um aus der an den Punkten r tatsächlich den Wert der Funktion zu erhalten, wurde die lokale Kraft auf die Gesamtkraft normiert. Obwohl die Graphen ähnlich sind, zeigt sich doch im Bereich kleiner Radien der erwähnte, und durchaus wichtige, Unterschied: Die durch Kraftabbildung gewonnenen Kurven liegen etwas unter den durch Nachstellung erhaltenen. Die Unterschiede erscheinen in dieser logarithmischen Darstellung gering, bedeuten aber viel.

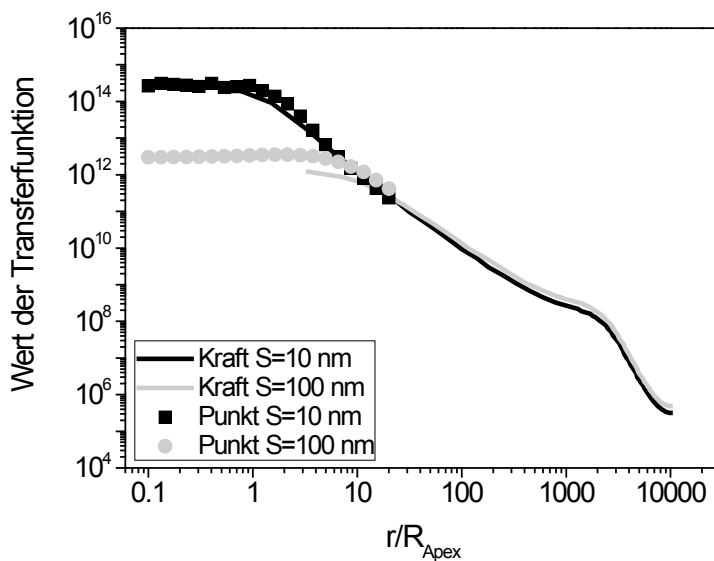


Abbildung 4.24: Vergleich von Kraftabbildung (Linien) und Nachstellung (Symbole) zur Gewinnung der Transferfunktion. Die für kleine Radien sichtbaren Unterschiede gestatten es nicht, die Methode zur Auflösungsbestimmung zu nutzen.

So wäre aus der Kraftabbildung nämlich ein deutlich geringerer Einfluss des Apex ablesbar, was sich nicht mit der analytischen Näherung und auch den Experimenten zur Auflösung vereinen lässt. Demnach kann diese, schnelle und einfache, Methode nur bedingt zu Aussagen herangezogen werden.

5. Ausgewählte experimentelle Ergebnisse

Neben der durchaus theorielastigen numerischen und analytischen Arbeit wurden Experimente an Proben aus verschiedensten Bereichen durchgeführt. Die interessantesten Ergebnisse dieser Experimente sollen hier vorgestellt werden, vor allem, um Theorie und Realität zusammenzubringen.

5.1 Allgemeines zu den Methoden – Typische Kontraste

Eine der ersten durchgeführten Messungen, zugleich eine der grundlegendsten, war die Bestimmung der SKM-Spannung für verschiedene Materialien. Dazu wurden Messungen mit einer TiN-beschichteten Spitze durchgeführt, die metallischen Proben waren dabei ebenfalls kommerziell beschichtete Cantilever (NtMdt). Der Fakt, dass sämtliche Metalle als dünne Filme (Dicke etwa 50 nm) auf Silizium vorlagen, mag einen Einfluss auf die Ergebnisse haben.

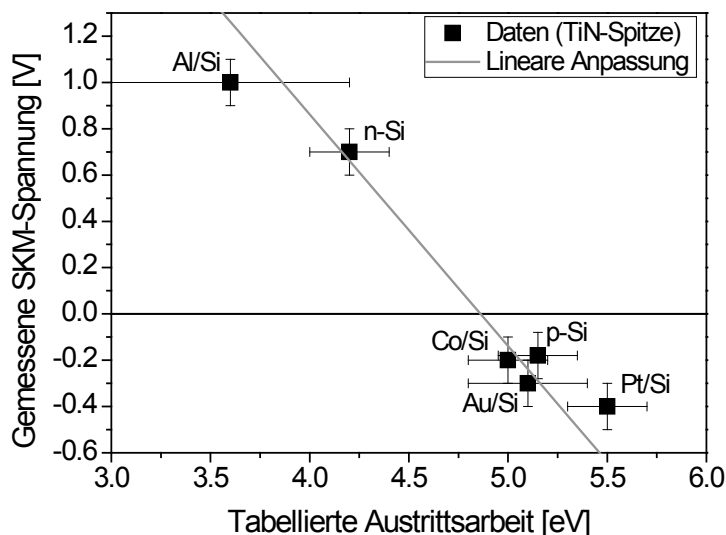


Abbildung 5.1: Gemessene SKM-Spannung für festes Spitzenmaterial und unterschiedliche Proben. Innerhalb der Fehlertoleranzen der Messung und der tabellierten Austrittsarbeiten ergibt sich ein linearer Zusammenhang.

Wie in Abbildung 5.1 illustriert, ergibt sich im Rahmen gewisser Fehlertoleranzen ein linearer Zusammenhang zwischen gemessener Spannung und Austrittsarbeit. Wie zu erwarten, misst demnach SKM die Differenz der Austrittsarbeiten von Spitze und Probe. Diese Erkenntnis ist allerdings nicht zu wertvoll, da aufgrund des fehlenden Wissens über die tatsächliche Austrittsarbeit der Spitze, welche im Übrigen durch das Auftreten unerwünschter zusätzlicher Spannungen überlagert sein kann, absolute Aussagen ohnehin schwierig sind.

Eine beispielhafte Abbildung, gewonnen aus einem SKM-Experiment an Aluminium auf Silizium, zeigt Abbildung 5.2. Offenbar korrelieren Topografie der Oberfläche und gemessenes Potential gut miteinander, von einigen Schwankungen abgesehen bleiben die Signale in den entsprechenden Bereichen stabil. Vergleicht man allerdings den theoretischen Verlauf, extrahiert aus einer zweidimensionalen Lösung der Poissongleichung ohne SKM-Simulation,

Ausgewählte experimentelle Ergebnisse

ergeben sich sowohl leichte Unterschiede in der absoluten Differenz, als auch in der Form des Profils.

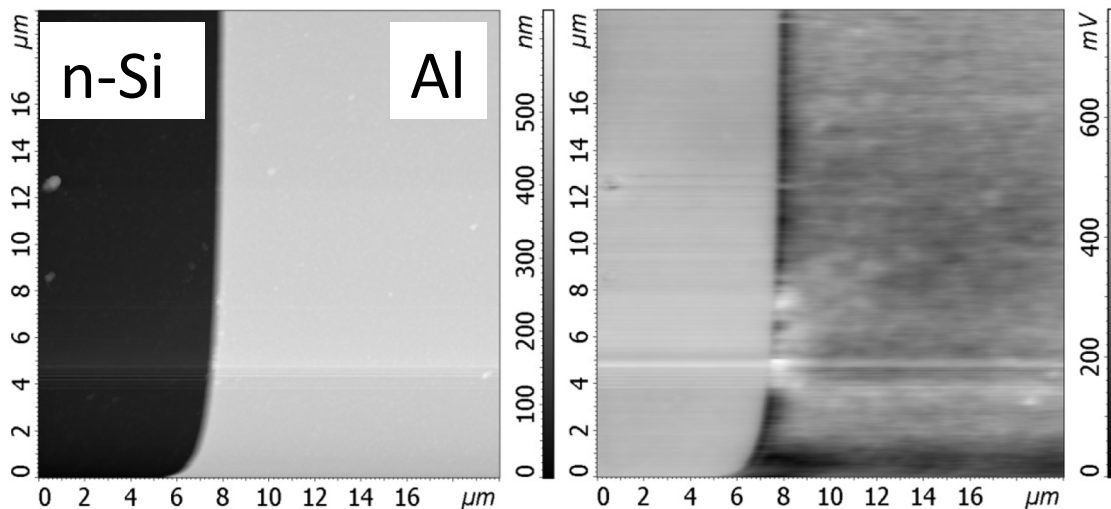


Abbildung 5.2: Topografie (links) und SKM-Karte (rechts) einer Aluminiumschicht auf Silizium. Augenscheinlich korrelieren Material und gemessene Spannung miteinander, was aufgrund der unterschiedlichen Austrittsarbeiten zu erwarten ist.

Die leichte Verringerung des absoluten Wertes um etwa 50 mV gegenüber der Rechnung mag an einem geladenen Oxid auf dem Aluminium liegen. Interessant ist das Auftreten einer starken Störung in der Nähe des Sprungs in der Topografie, was wahrscheinlich durch den hohen lateralen Einfluss der Stufe zu erklären ist.

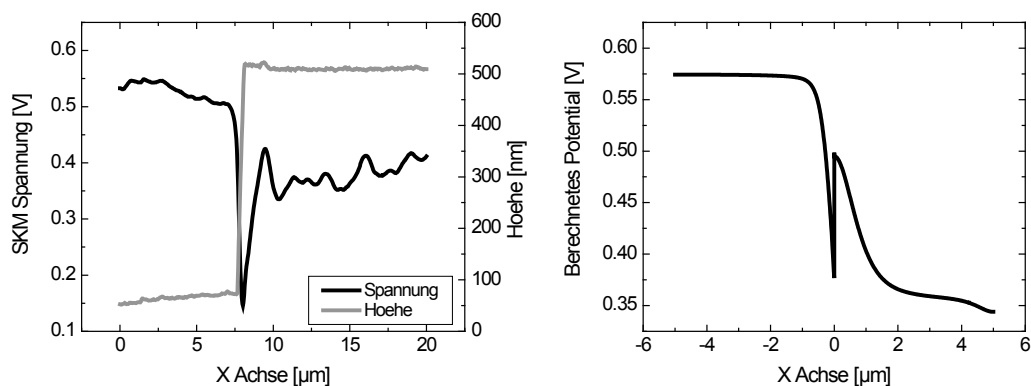


Abbildung 5.3: Vergleich des gemessenen SKM-Profiles (links, schwarze Spur) mit dem mittels 2D-Simulation abgeschätzten Verlauf (rechts). Im Großen und Ganzen sind die Kurven als sehr ähnlich zu beschreiben, Unterschiede erklären sich durch möglicherweise geladene Oxidschichten und dem doch relativ großen Sprung in der Probenhöhe.

Auch die SCM-Methode liefert einen guten Kontrast, wie Abbildung 5.4a illustriert. Der Unterschied in den gemessenen Amplituden für den Silizium- und Aluminiumbereich ergeben sich dabei durch kapazitiv unterschiedliche Oxidschichten. Interessant ist der deutlich bemerkbare Einfluss der Topografiestufe, welcher sich in starken Überschwingern ausdrückt (Abbildung 5.4b).

Das vorgestellte, doch sehr einfache Beispiel, sollte lediglich typische Kontrastmechanismen verdeutlichen und auf mögliche experimentelle Probleme hinweisen. Das mit Abstand größte Problem hierbei stellen offenbar starke Schwankungen im Höhenprofil dar, ein Umstand der glücklicherweise bei üblichen Proben nicht vorliegt und deshalb an dieser Stelle nicht tiefgehend besprochen werden soll.

Nicht präsentiert werden sollen in diesem Kapitel zum Beispiel Ergebnisse von Experimenten an dünnen dielektrischen Schichten, die mittels Pulsed Laser Deposition auf Silizium abgeschieden wurden. Die Methoden konnten dabei erfolgreich genutzt werden, um sowohl Ladungszustand als auch kapazitive Eigenschaften der Schichten zu charakterisieren [31]. Ebenfalls wurde vor allem SKM genutzt, um auftretende Kontaktpotentiale in organischen Halbleitern zu bestimmen, wobei die Strukturen unter äußerer Vorspannung vermessen wurden [32]. Die mit möglichen Phasenumwandlungen und Segregation verbundenen Potentialdifferenzen, bestimmt an verschiedenen siliziumbasierten Proben, wurden in [33] und [34] veröffentlicht. Alle diese Ergebnisse unterstreichen die vielseitige Anwendbarkeit der Methode, vor allem in Kopplung mit der Möglichkeit simultan zu Topografie und Potential zum Beispiel die Härte der Oberfläche zu beurteilen [35].

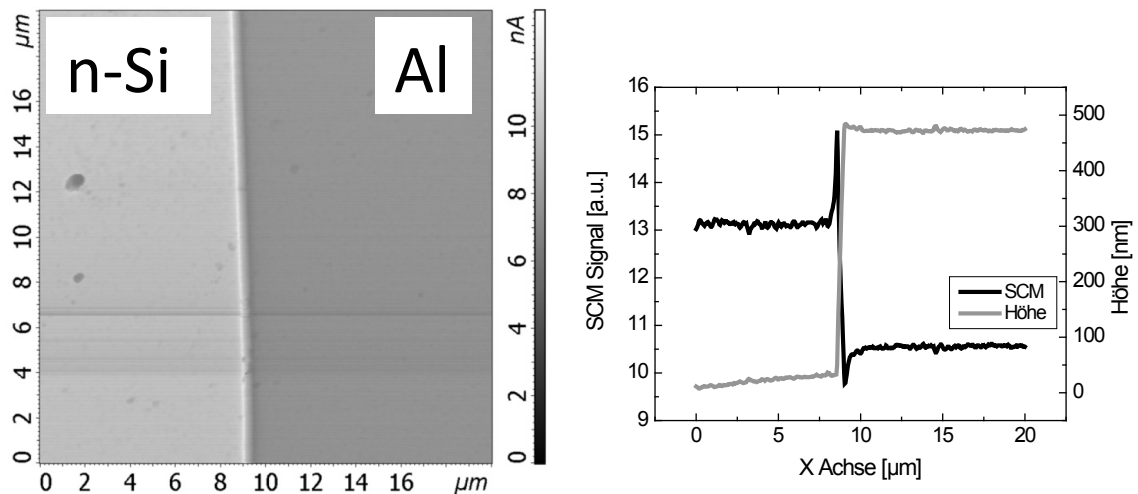


Abbildung 5.4: (Links) SCM-Karte der Beispielprobe. Der recht deutliche Kontrast ist in kapazitiven Unterschieden, wahrscheinlich von vorhandenen Oxidschichten, begründet. Ein extrahiertes Profil (rechts) zeigt wiederum deutlichen Einfluss der Kante in der Probenhöhe.

Weitaus interessanter als die genannten Beispiele erscheinen zwei Klassen von Proben, die im Weiteren genauer dargestellt werden. Anhand dieser lässt sich nämlich das Potential der kraftbasierten Methoden gut erläutern, vor allem in Hinblick auf häufig auftretende Probleme.

5.2 Gebondete Silizium-Wafer

Die ursprüngliche Motivation zur Herstellung solcher Strukturen lag in der erhofften Möglichkeit, reguläre Variationen des Oberflächenpotentials auf nm-Skala zu erzeugen. Diese Variationen sollten dann zur strukturierten Anlagerung von Biomolekülen genutzt werden, was wiederum mögliche Anwendung in Sensor-Szenarien finden kann.

Ausgewählte experimentelle Ergebnisse

Eine genauere Beschreibung der Herstellung der Proben entnimmt man [36], hier soll auf detaillierte Beschreibung verzichtet werden. Für die Charakterisierung mittels AFM wurden Schnitte der Proben hergestellt und mechanisch poliert. Des Weiteren wurden an den Seiten der Probe Kontakte zum Anlegen einer Vorspannung angebracht (Abbildung 5.5 links). Die untersuchte Oberfläche weist einige Spuren des Polierens auf, ist im Großen und Ganzen aber glatt und zeigt keine Spur des Interface (Abbildung 5.5 rechts).

Alternativ durchgeführte Messungen mittels der EBIC-Methode (Electron Beam Induced Current), welche prinzipiell eine Abbildung der lokalen Diffusionslänge erlaubt, ergaben eine deutliche Signatur des Interface, wahrscheinlich verursacht durch eine sehr hohe Rekombinationsrate. Auch konnte aus diesen Messungen, zusammen mit einigen theoretischen Betrachtungen, ein wirkendes Potential in der Größenordnung von 100 mV am Übergang zwischen den Wafern bestimmt werden [37].

Zunächst wurde senkrecht zum Interface die Verteilung des Oberflächenpotentials mittels SKM gemessen. Die Nutzung des zweiphasigen Setups erlaubt dabei die Justage des Arbeitsabstandes im zweiten Durchgang, was wie beschrieben (Kapitel 3) eine Entfaltung zulässt. Die in Abbildung 5.6a dargestellten SKM-Profile für zwei verschiedene Höhen zeigen dabei zwei wichtige Punkte: Zum einen ist tatsächlich an der Position des Interface ein deutlicher Unterschied im Oberflächenpotential verglichen mit dem Rest der Probe erkennbar und zum anderen hängt die messbare Potentialdifferenz merklich von der Höhe ab.

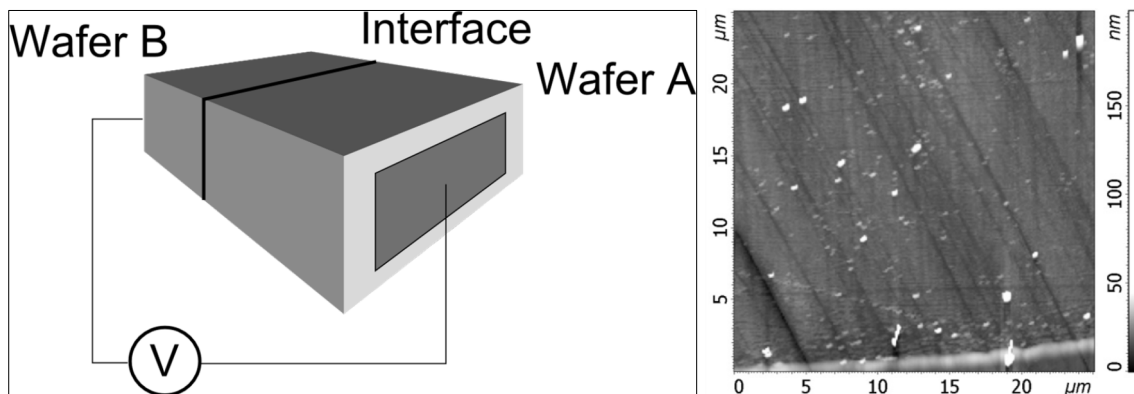


Abbildung 5.5: (Links) Schematische Skizze der Probe. Es wurden Schnitte quer zum Interface zwischen den beiden Wafern A und B hergestellt. Zusätzliche Kontakte erlauben das Anlegen einer äußeren Spannung V. Die im rechten Bild dargestellte Topographie der Oberfläche zeigt wenige Spuren der Präparation aber keine Signatur des Interface.

Während es durch Extrapolation möglich ist, Potentialdifferenzen im Bereich von 100 mV zu deklarieren, ergibt sich ein prinzipielles Problem bei der Interpretation der Ergebnisse. In der zugrunde liegenden Theorie [36] geht man für die p-dotierten Wafer von positiven Linienladungen entlang der individuellen Versetzungslinien aus. Nutzt man die dort publizierten Werte für die Linienladungsdichte unter Annahme eines radialen Verarmungsmodells [38], ergeben sich die genannten Werte für die erwartete Potentialdifferenz, allerdings müsste diese positiv und nicht negativ wie in SKM gemessen sein. Die Weiten der Verarmungszonen aus dem Modell, welche die Weite der Potentialänderung beschreiben, sind mit etwa 1 µm geringer als die gemessenen Weiten von einigen µm (siehe auch Anhang A)

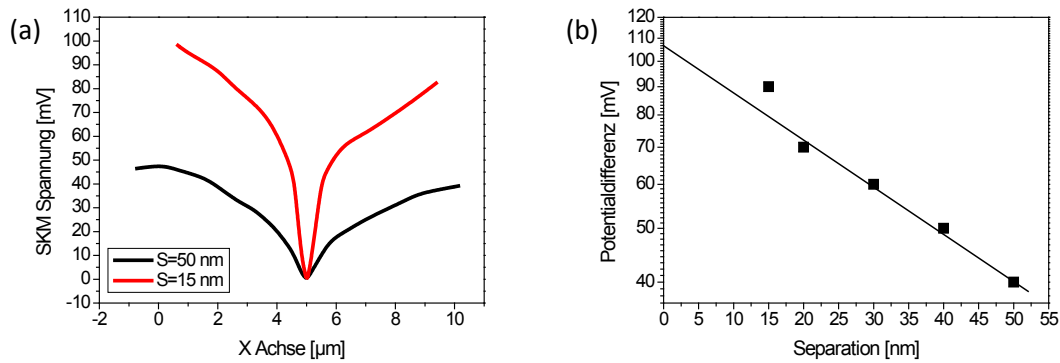


Abbildung 5.6: Typische SKM-Profil gemessen senkrecht zum Interface zweier gebondeter Wafer (links) und gemessene Potentialdifferenz in Abhängigkeit der Separation S zwischen Spitze und Probe. Obwohl die Form der Profile nicht überrascht, der extrapolierte Wert für die Differenz richtig erscheint, erzeugt die Messung doch eine grundsätzlich falsch erscheinende Aussage.

Nimmt man an, dass zumindest die Weiten ein Artefakt der einfachen Modellrechnung sind, kann man nun durch korrekte Berechnung Vergleichswerte erzeugen. Dazu wird ein zweidimensionales Modell, unter Berücksichtigung der Symmetrie am Interface, erstellt, in welchem die Poissongleichung gelöst wird (Abbildung 5.7a). Offenbar ergeben sich an der Oberfläche Potentialverläufe die der Form nach mit den gemessenen vergleichbar sind, nur hat die Änderung das falsche Vorzeichen. Die Weiten sind, wie vom Modell vorhergesagt, nicht vergleichbar. An diesem Punkt könnte man wohlwollend annehmen, dass die gemessenen breiteren Ausdehnungen, vor allem für große Separationen, lediglich messbedingt sind. Da allerdings eine negative Aufladung der Versetzungen unwahrscheinlich ist, muss die Ursache für die Beobachtungen eine andere sein.

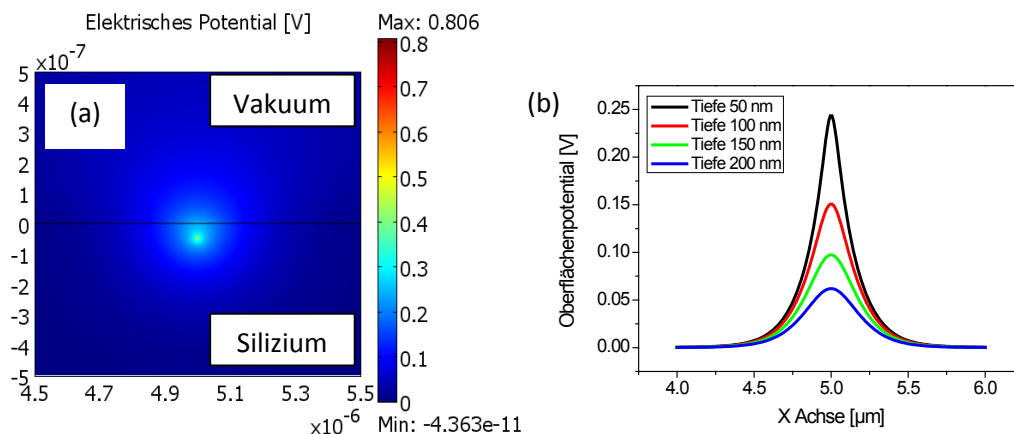


Abbildung 5.7: FEM-Lösung des Poissonproblems mit Linienladung in Silizium (a) und extrahierte Profile des Oberflächenpotentials für verschiedenen Tiefen der Linienladung unterhalb der Oberfläche (b). Die so generierten Verläufe ähneln den gemessenen SKM-Profilen, weisen allerdings das falsche Vorzeichen auf.

Um genauere Aussagen vornehmen zu können, wurden vorspannungsabhängige Messungen durchgeführt. Dazu wurde an den senkrecht zum Interface liegenden Kontakten während des Scans Spannungen angelegt. Um Vergleichbarkeit zu erzeugen, wurden diese Spannungen immer für einige Linienscans konstant gehalten und dann geändert. Die resultierende SKM-Abbildung (Abbildung 5.8a) zeigt interessante Effekte: Zum einen erscheint das Potential in den

Ausgewählte experimentelle Ergebnisse

beiden Wafern (oberer und unterer Bereich des Bildes) glatt, also quasi feldfrei, und die Wirkung des Interface verschwindet nicht. Würde man von einfacher Aufladung nebst Verarmung ausgehen, müsste ein äußeres Potential in der gleichen Größenordnung wie das in der Probe vorhandene dieses maskieren. Dieser Effekt ist nicht zu beobachten. Extrahiert man Profile (Abbildung 5.8b), verdeutlichen sich diese Beobachtungen.

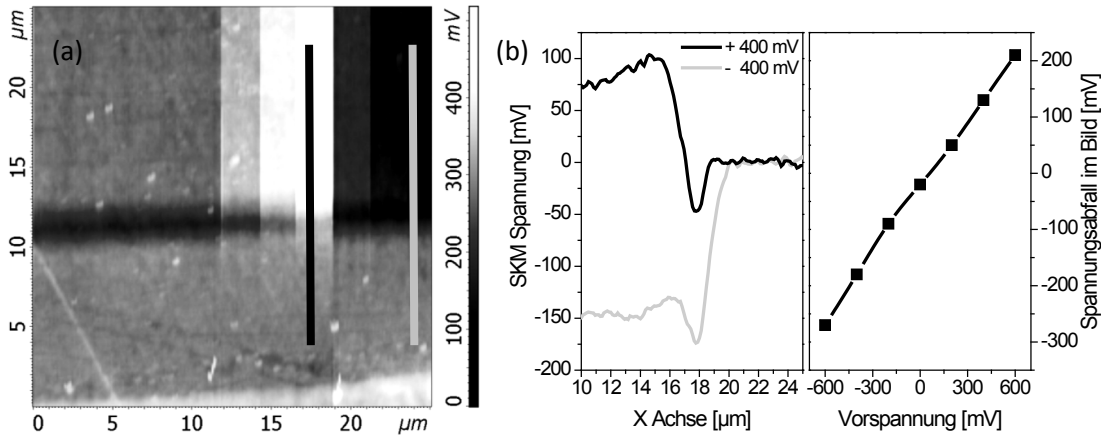


Abbildung 5.8: SKM-Karte erzeugt mit variierender Vorspannung zwischen -0.5 V und 0.5 V (a) und extrahierte Profile entlang der gekennzeichneten Linien (b). Zusätzlich ist der Spannungsabfall im abgebildeten Bereich dargestellt, der etwa 50% der angelegten Spannung beträgt. Offenbar gelingt es nicht, die Wirkung des Interface zu maskieren, was zusammen mit dem hohen senkrecht wirkenden Widerstand eine hohe Rekombination suggeriert.

Stellt man den Spannungsabfall im Scanbereich über der äußeren Spannung dar, ergibt sich, dass in der Nähe des Interface etwa die Hälfte der äußeren Spannung abfällt. Damit kann man, recht vorsichtig, effektive Ladungsdichten abschätzen:

$$1 = \frac{U_W}{U_I} = \frac{R_W}{R_I} \approx \frac{\frac{L_W}{q\mu N_A}}{\frac{L_I}{q\mu n}} = \frac{L_W}{L_I} \frac{n}{N_A} \quad (5.2.1)$$

Dafür wurde angesetzt, dass die Spannungsabfälle in Wafer (Index W) und am Interface (Index I) sich wie die entsprechenden Widerstände verhalten, wobei diese dann über die Ladungsdichten und die Ausdehnungen L definiert wurden. Da sich nun diese Ausdehnungen etwa um einen Faktor von 1000 unterscheiden, muss die effektive Ladungsdichte n in der Nähe des Interface um diesen Faktor verringert gegenüber der Dotierungsdichte N_A sein, und zwar unabhängig von der äußeren Spannung.

Der Meinung des Autors nach weist das auf ein rekombinationsbestimmtes Phänomen hin. Der Form nach ähnliche Verläufe des Oberflächenpotentials unter Beleuchtung wurden im Übrigen auch an Dioden gemessen [39], die Autoren argumentieren dort, dass der laterale Verlauf des Potentials senkrecht zum p-n-Übergang durch die diffusionsbestimmte Ladungsträgerkonzentration gegeben ist. Das dort genutzte eindimensionale Modell beschreibt den Verlauf der Konzentration durch eine Lösung der Diffusionsgleichung mit entsprechenden Randbedingungen unter der Annahme konstanter Generation. Da die eindimensionalen Lösungen dieser Gleichungen im Anhang C besprochen werden, soll an dieser Stelle nur die grundlegende Ortsabhängigkeit angegeben werden:

$$n(x) = n_0 + G\tau \left(1 - \frac{s}{s + \frac{D}{L}} \exp\left(-\frac{x}{L}\right) \right) \quad (5.2.2)$$

In Gleichung (5.2.2) sind offenbar die bestimmenden Parameter die Diffusionslänge L und die Interfacerekombinationsgeschwindigkeit s , die Bedeutung der anderen Parameter entnimmt man dem genannten Anhang. Offenbar stellt sich, abhängig von den Rekombinationsparametern, ein exponentieller Verlauf der Ladungsträgerdichte ein. Setzt man nun diesen in (A3.16) ein, ergibt sich folgender erwarteter Verlauf für das Potential:

$$V(x) = V_i \ln \left(1 + \frac{G\tau}{n_0} \left(1 - \frac{s}{s + \frac{D}{L}} \exp\left(-\frac{x}{L}\right) \right) \right) \quad (5.2.3)$$

Der angegebene laterale Verlauf trägt damit sowohl die Signatur der Diffusionslänge als auch der Rekombinationsgeschwindigkeit. Schwierig scheint im Moment allerdings eine sinnvolle Definition dieser Diffusionslänge, da im Gegensatz zu der zitierten Quelle ein eindimensionaler Ansatz im Allgemeinen nicht genügen dürfte. Deshalb wurde, um wiederum Abschätzungen zu erhalten, die Diffusionsgleichung modellhaft in zwei Dimensionen mit Flussrandbedingungen (Oberflächenrekombination) an $x=0$ und $y=0$ gelöst. Die komplette Verteilung (Abbildung 5.9 links) der Konzentration kann nicht analytisch angegeben werden, allerdings zeigen die an der Oberfläche $y=0$ extrahierten Verläufe exponentielle Abhängigkeit von der Koordinate x (Abbildung 5.9 rechts). Offensichtlich ist ein starker Einfluss der Oberflächenaktivität, definiert durch eine Rekombinationsgeschwindigkeit s_0 . Um nutzbare Aussagen zu gewinnen, wurden parametrisch Lösungen in Abhängigkeit dieses Parameters erzeugt und die entsprechenden Kurven exponentiell angepasst.

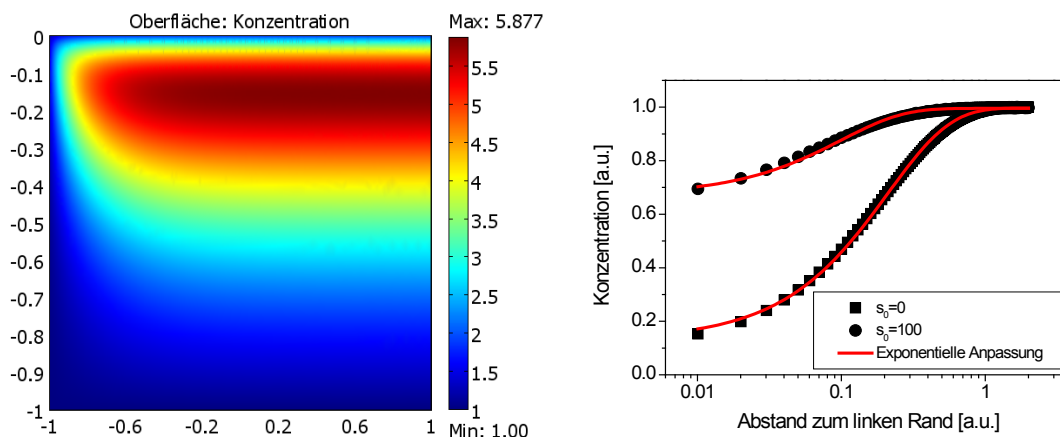


Abbildung 5.9: (Links) Typische numerische Lösung der Diffusionsgleichung mit Flussrandbedingungen in x- und y-Richtung. Die Verteilung der Konzentration ist kompliziert und kann nicht analytisch beschrieben werden. Die dargestellten Verläufe der Konzentration an der Oberfläche ($y=0$) zeigen exponentielles Verhalten mit Abklingkonstanten, die von der Rekombinationsgeschwindigkeit an dieser Oberfläche abhängen (rechts).

Den Verlauf der wichtigen Anpassungsvariablen, Amplitude und Abklingkonstante, illustriert Abbildung 5.10. Offenbar sinken sowohl die Amplitude, somit die maximale

Ausgewählte experimentelle Ergebnisse

Ladungsträgerdichte, als auch die Abklingkonstante, die effektive Diffusionslänge an der Oberfläche, nichtlinear mit der Rekombinationsgeschwindigkeit der Oberfläche. Der Verlauf lässt sich, an dieser Stelle lediglich empirisch motiviert, gut durch eine rationale Funktion anpassen. Überträgt man diese Ergebnisse, ist also zu erwarten, dass sich abhängig von der Aktivität der Oberfläche senkrecht zum Bonding-Interface unter Beleuchtung exponentielles Wachstum der Ladungsträgerkonzentration einstellt. Geht man von nahezu unendlich großer Rekombinationsrate an diesem Interface aus, was aufgrund der hohen Defektdichte gerechtfertigt scheint, kann vermutet werden, dass in diesem Fall nicht die Potentialverteilung bedingt durch Ladungsanhäufung sondern vielmehr eine Photospannung detektiert wird. Diese ist nun ein Maß für die Ladungsträgerdichte am Ort der Messung und vermag so, effektive Lebensdauern abzubilden. Die Größenordnung für die Differenz der Photospannungen in der Nähe des Interface und weit entfernt davon liegt bei etwa 100 mV, wenn man von den Parametern des im AFM-System verbauten Detektionslasers ausgeht. Dieser Schätzwert entspricht dem gemessenen, was die Gültigkeit der Annahmen und Modellierung unterstützt.

Zusätzlich zu den relativ einfachen Modellierungen der Diffusionsproblematik wurden die vollständigen Drift-Diffusions-Gleichungen implementiert in zwei Dimensionen, um sowohl verlässlichere Werte für die Photospannung zu erhalten als auch das Verhalten der Probe unter äußerer Spannung zu simulieren. Wichtige Resultate dieses Vorgehens zeigt Abbildung 5.11. Im linken Bild dargestellt ist der modellierte Verlauf der Potentialdifferenz zwischen Interface und Orten weit entfernt davon in Abhängigkeit der Lichtintensität (normiert auf den erwarteten Wert im genutzten System). Deutlich ersichtlich ist auch hier der Einfluss der Rekombinationsgeschwindigkeit der Oberfläche, für quasi inaktive Flächen ergibt sich demnach der gemessene Wert von etwa 100 mV. Vergleicht man zusätzlich die vorspannungsabhängigen Verläufe im rechten Bild der Abbildung mit denen in Abbildung 5.6 und Abbildung 5.8 erkennt man durchaus Vergleichbarkeit.

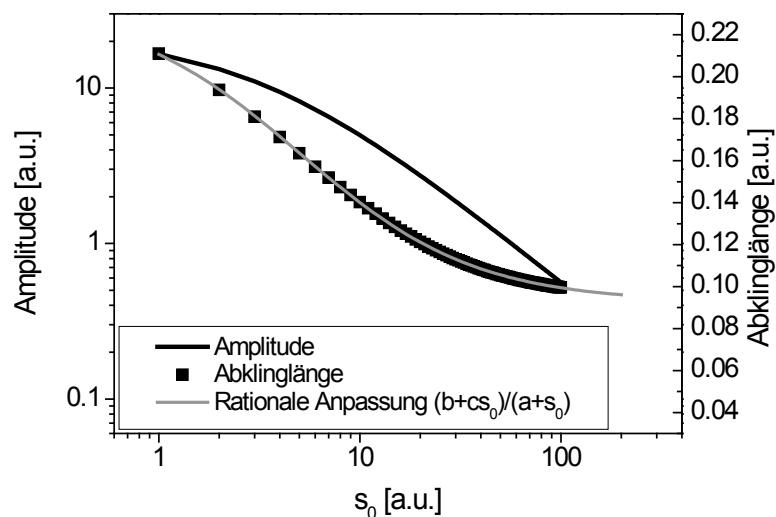


Abbildung 5.10: Verlauf der Amplitude der Konzentration und der effektiven Diffusionslänge von der Oberflächenrekombinationsrate für zweidimensionale Modellierung der Diffusionsgleichung. Beide Parameter skalieren nichtlinear, was für Messungen des Oberflächenpotentials starke Abhängigkeit von Wert und lateraler Ausdehnung bedeutet.

Tatsächlich ergeben sich Halbwertsbreiten des Potentialminimums in der Größenordnung einiger μm , was sowohl den gemessenen Werten als auch typischen Diffusionslängen in Silizium entspricht (einige Werte der Volumendiffusionslänge entnimmt man [40]). Allerdings bleibt darauf hinzuweisen, dass die Modelle nicht vollständig sind, zum Beispiel wurde der Effekt der Aufladung des Interface unterdrückt. Zusätzlich mögen die einfachen Annahmen für die Randbedingungen nicht ausreichend sein, sie erlauben allerdings eine numerische Behandlung mit akzeptablen Anforderungen an die genutzte Hardware.

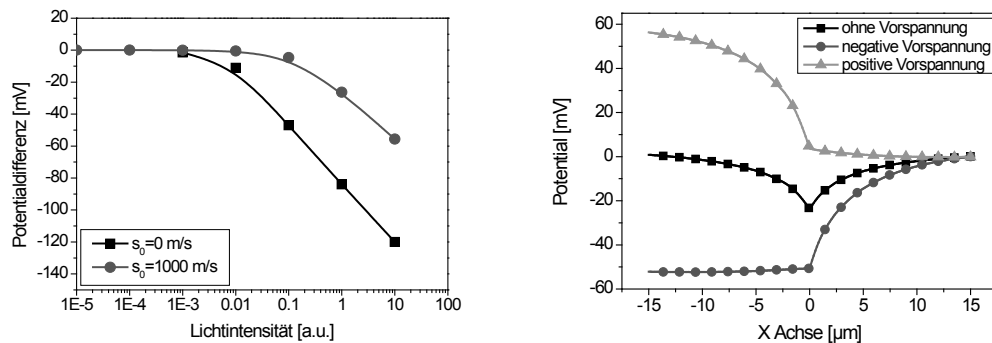


Abbildung 5.11: Berechnete Potentialdifferenz in Abhängigkeit der Beleuchtung (links) und modellierte Verläufe des Potentials unter Vorspannung (rechts). Die Potentialdifferenzen variieren stark mit der Aktivität der Oberfläche sind aber unter vernünftigen Annahmen im gemessenen Bereich anzusiedeln. Die typische Form der Kurven und auch die Halbwertsbreite des Potentialminimums im rechten Bild sind vergleichbar mit gemessenen Werten, was die Annahmen unterstützt.

Neben den beschriebenen SKM-Experimenten wurde auch die nc-SCM-Methode auf die Proben angewendet. Eine typische erhaltene Karte zeigt Abbildung 5.12 (links). Deutlich erkennbar ist, neben einigen dunkel erscheinenden Störungen der Oberfläche (markante Punkte können mit Abbildung 5.5 verglichen werden), das Interface als Linie verringerten Signals. Ein senkrecht zu dieser ausgeprägten Linie extrahiertes Profil zeigt das rechte Bild der Abbildung. Die Strukturierung dieser Spur ist deutlich komplexer als die der SKM-Profile, eine Interpretation erfordert demnach größeren Aufwand.

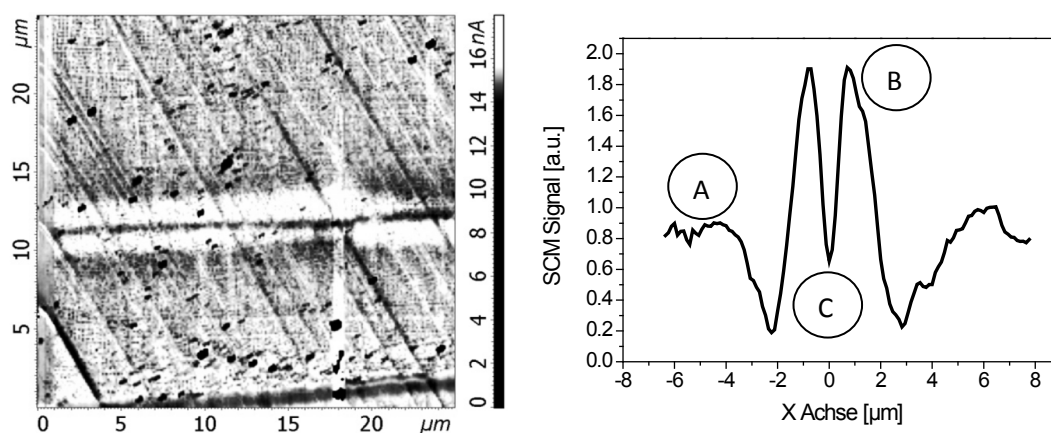


Abbildung 5.12: SCM-Karte (links) und extrahiertes Profil senkrecht zum Interface (rechts). Auch in dieser Abbildung hinterlässt die Kontaktstelle der Wafer eine deutliche Signatur, allerdings ist der komplexe Verlauf des Profils schwierig zu erläutern. Die Buchstaben dienen der Identifizierung prägnanter Bereiche.

Die prinzipielle Abhängigkeit des SCM-Signals von Probenparametern wurde in Kapitel 3 beschrieben. Auf spezielle Effekte für halbleitende Proben wurde in Kapitel 3.5 eingegangen. Die dort präsentierten Ergebnisse für virtuelle Scans senkrecht zu p-n-Übergängen zeigen den Einfluss von Potentialdifferenzen auf die kapazitive Antwortfunktion. So verschieben sich die Signal-Spannungs-Kurven proportional mit dem lokalen Potential, was bei fester Spannung an der Spitze des AFM-Systems ganz unterschiedliche Kontraste erzeugen kann. Ein weiterer Einfluss ist natürlich die lokale Ladungsträgerdichte, welche maßgeblich die Oberflächenkapazität und damit den Messwert bestimmt. Unglücklicherweise überlagern sich diese beiden Effekte, weshalb die im Folgenden angebotene Deutung der Ergebnisse stark phänomenologischen Charakter ohne Anspruch auf vollständige theoretische Hinterlegung hat.

Geht man zunächst davon aus, dass für Messspannungen in der Nähe des Oberflächenpotentials, was für die präsentierte Messung der Fall ist, tatsächlich die Probenkapazität das Signal bestimmt (siehe Gleichung 3.4.3), können Argumentationen über Oberflächenkapazitäten von Halbleitern zur Interpretation genutzt werden. Dazu können nun drei signifikant unterschiedliche Bereiche der Probe identifiziert werden: Weit entfernt vom Interface ist das Signal konstant (A), direkt am Interface deutlich abgesenkt (C) und in der Nähe davon stark überhöht (B). Nimmt man nun an, dass die typische Form der Antwortfunktion der Kapazitäts-Spannungs-Kurve eines MOS-Systems folgt (siehe zum Beispiel [23]), ergibt sich folgendes Bild: Die Bereiche „A“ verhalten sich „normal“, das heißt, die Kapazität sollte aufgrund der Messfrequenz und der relativ hohen Minoritätsladungsträgerdichte einer Niederfrequenzkurve ähneln. Die Signalerhöhung in den Bereichen „B“ kann durch Verschiebung der Spannungsachse der Signal-Kurven beschrieben werden, die gemessene negative Spannung auf der Probenoberfläche verschiebt die Kurven in Richtung höherer Spannung, was einen Anstieg der Kapazität und damit des Signals bewirkt. In der Nähe des Interface wird dieser Effekt durch einen zweiten überlagert, nämlich die Absenkung der effektiven Ladungsträgerdichte. Tatsächlich kann man den umgekehrten Effekt an realen MOS-Systemen messen, Beleuchtung führt hier zur Anhebung der Kapazität, was im Prinzip der Wandlung von dynamischer (Hochfrequenz) zu statischer (Niederfrequenz) Antwort entspricht [41]. Ein mögliches Szenario unter Zuhilfenahme dieser Erklärung illustriert Abbildung 5.13: Ausgehend von Punkten mit der in schwarz skizzierten Kapazitäts-Spannungs-Kurve gelangt man zu den Punkten nach rechts verschobener Kennlinie (grün), welche demnach höhere Kapazität bei derselben Spannung zeigen. An anderen Punkten, an denen die Form der Kurve verändert ist (rot) erhält man wiederum ein geringeres Signal. Ähnliche Effekte ergeben sich im Übrigen auch in „Standard“ SCM-Experimenten, das Problem ist also genereller Natur [42].

Der Unterschied zwischen den SKM-Profilen und den in SCM ermittelten (Abbildung 5.13 rechts) wird somit vor allem durch die unterschiedlichen charakteristischen Längen von Ladungsträgerdichten und Photospannung erzeugt. Während die Kapazität, und damit auch das SCM-Signal, etwa direkte Proportionalität zur Ladungsträgerdichte zeigen sollte, ist die Verknüpfung des SKM-Signals mit dieser Dichte funktional komplizierter. In Konsequenz folgt somit der kapazitive Messwert der Dichteverteilung, mit der charakteristischen Länge gleich der Diffusionslänge, und die Oberflächenspannung ist durch eine andere Funktion mit geringer Abklingkonstante bestimmt. Davon ausgehend erkennt man in den SCM-Profilen zunächst das Absinken der Ladungsträgerdichte ohne Überlagerung von Spannungen, dann den Anstieg des

Signals aufgrund der Verschiebung der Antwortfunktion und schließlich wieder das Absinken aufgrund der völlig anderen Form der Antwort.

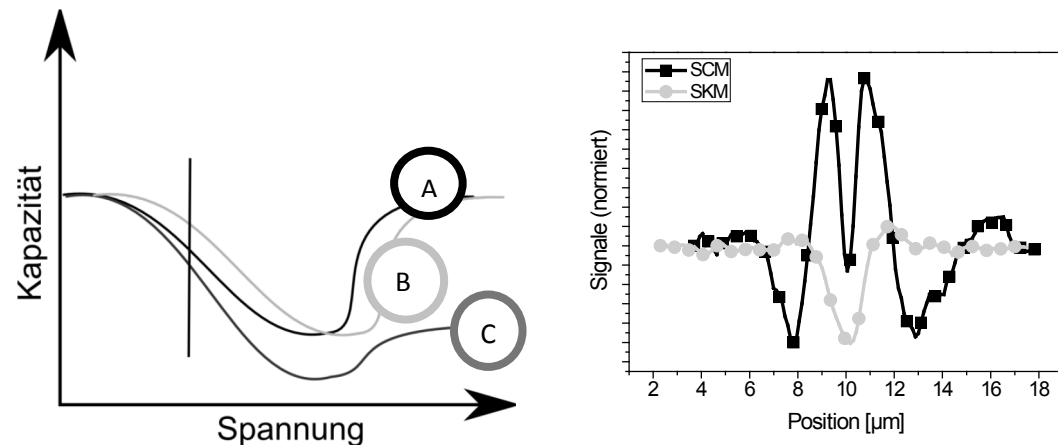


Abbildung 5.13: (links) Einfaches Modell zur Beschreibung der SCM-Ergebnisse der gebondeten Wafer. Es kann davon ausgegangen werden, dass zum Einen das vorhandene Potential die Antwortkurve verschiebt (A-B) und auch die lokale Ladungsträgerdichte einen Einfluss zeigt (B-C). Wird eine Messung mit der durch die Linie gekennzeichneten Spannung durchgeführt, ergeben sich qualitativ die Ergebnisse. Der rechts dargestellte Vergleich von SKM und SCM verdeutlicht die unterschiedlichen Längenskalen.

Obwohl sich mit diesem einfachen Modell die Ergebnisse erklären lassen, muss darauf hingewiesen werden, dass die Ausführungen hochgradig spekulativ sind. Eine gute analytische Beschreibung beziehungsweise numerische Modellierung muss der Autor an dieser Stelle schuldig bleiben.

5.3 Dotiergitter in Silizium

Eine andere Probenklasse mit noch größerem Potential für die Entwicklung der Methodik sind Dotiergitter in Silizium. Derartige Proben wurden durch Standard-CMOS-Prozessierung am IHP (Frankfurt/Oder) hergestellt. Eine genauere Erläuterung der Präparation entnimmt man [43], hier soll lediglich die grundlegende Struktur der untersuchten Proben erläutert werden.

Durch die Implantation von Arsenionen mit fester Energie (45 keV) und verschiedenen Dosen durch eine Maske hindurch wurden im p-dotierten Substrat (Dotierdichte $N_A \approx 10^{15} \text{ cm}^{-3}$) periodische p-n-Übergänge mit Gitterkonstanten von 360 nm eingebracht.

Um die elektronische Struktur abzuschätzen, wurde zunächst der Implantationsschritt mittels „SRIM“ [44] nachgestellt, um die Verteilung der Ionen bestimmen zu können. Das Ergebnis für die Tiefenverteilung, eine zusätzlich eingefügte Schicht auf der Probe wurde berücksichtigt, zeigt Abbildung 5.14 (links). Wie zu erwarten, ergibt sich ein Gaußprofil für die Ionendichte, die Breite beträgt etwa 25 nm. Die laterale Verteilungsbreite unter einer Maske kann ebenfalls mit genannter Software ermittelt werden, es ergaben sich Werte um 10 nm. Mit dem Wissen um diese Parameter kann nun die erwartete Dotierdichte als Produkt von Gaußverteilung in Tiefe und Fehlerfunktion in Breite modelliert werden. Eine so erhaltene Dichtefunktion zeigt Abbildung 5.14 (rechts).

Ausgewählte experimentelle Ergebnisse

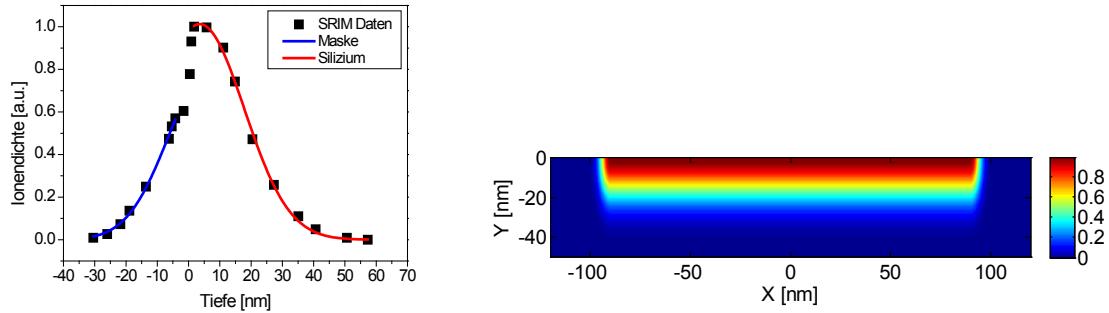


Abbildung 5.14: Tiefenverteilung (links) und zweidimensionale Verteilung der implantierten Ionen (rechts), berechnet mit Hilfe von „SRIM“. In der Tiefe ergibt sich ein Gaußprofil, lateral ist es eine Fehlerfunktion. Die Stufe in der Tiefenverteilung entsteht durch das Aufbringen einer zusätzlichen Schicht, welche das Maximum der Dotierdichte an der Oberfläche gewährleistet.

Die erhaltene Dichtefunktion wurde dann genutzt, um die Potentialverteilung numerisch zu berechnen. Dazu wurden einige der, mit der Dotierungsdichte N_D skalierten, Dichtefunktionen für die Implantationsdichte angesetzt und die Poissonsgleichung (A.6) mit angepasstem Referenzpotential gemäß Gleichung (5.3.1) gelöst.

$$\Delta V = \frac{2qn_i}{\varepsilon} \left(\sinh \left(\frac{V - V_t \cdot \operatorname{asinh} \left(\frac{N_D(x,y) - N_A}{2 \cdot n_i} \right)}{V_t} \right) + \sinh \left(\frac{V_F}{V_t} \right) \right) \quad (5.3.1)$$

Eine typische numerische Lösung ist in Abbildung 5.15 (links) dargestellt, der Verlauf des Potentials auf der Oberfläche wird in Abbildung 5.15 (rechts) für verschiedene Dichten N_D präsentiert.

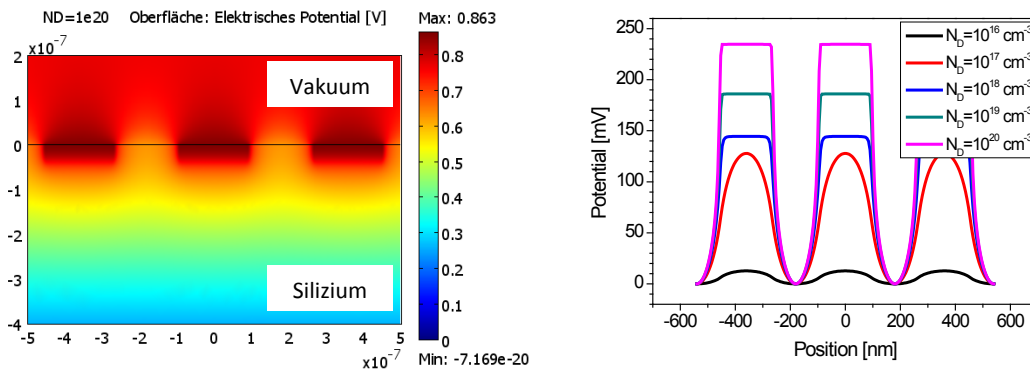


Abbildung 5.15: Abbildung der numerischen Lösung für Dotiergitter (links) und extrahierter Verlauf des Oberflächenpotentials für verschiedene Dotierungsdichten (rechts). Im geplanten Bereich der Dichten ist das Potential innerhalb der implantierten Bereiche quasi konstant, außerhalb davon erscheint der Verlauf einer Kettenlinie zu gleichen.

Offenbar gibt es einen starken Einfluss der Dotierung sowohl auf Form als auch auf Wert der Funktion. Ein wichtiger, weil messbarer, Parameter sollte die Variation des Potentials auf der Oberfläche sein. Dieser ist exemplarisch für einige Dichten extrahiert worden und in Abbildung

5.16 dargestellt. Man erwartet demnach, ausgehend von den in der Herstellung angestrebten Implantationsdichten zwischen 10^{17} und 10^{20} cm^{-3} Variationen in der Größenordnung von einigen 100 mV. Die durch Linien angegebenen Näherungsfunktionen sind analytisch schwer herzuleiten, im Fall recht großer Dichten N_D darf aber davon ausgegangen werden, dass der p-dotierte Bereich nahe der Oberfläche in Inversion ist. Demnach ist das Potential in diesem Bereich nur noch schwach von äußeren Schwankungen beeinflusst, es stellt sich als gemessene Variation ΔV die Differenz der Fermi-niveaus der implantierten und invertierten Bereiche ein:

$$q\Delta V = \left(E_i + kT \ln \left(\frac{N_D}{n_i} \right) \right) - \left(E_i + kT \ln \left(\frac{N_A}{n_i} \right) \right) = kT \ln \left(\frac{N_D}{N_A} \right) \quad (5.3.2)$$

Der Grad der Näherung ist nicht hervorragend, liegen doch die so geschätzten Werte etwa 20 % zu hoch. Für die weiteren Betrachtungen spielt diese Diskrepanz aber keine Rolle, für Vergleiche werden ausschließlich die numerisch gewonnenen Werte herangezogen.

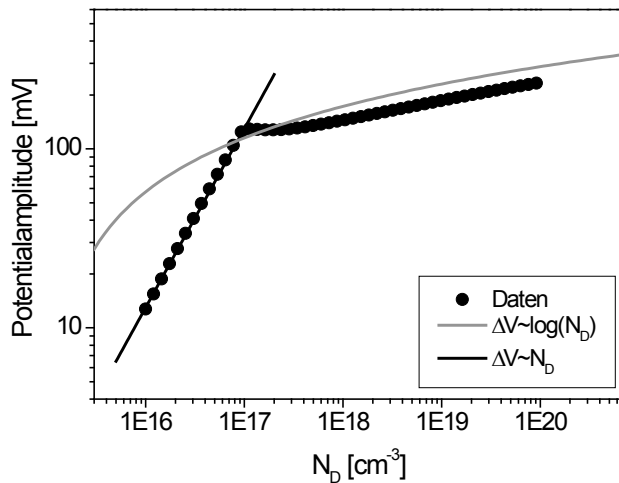


Abbildung 5.16: Abhängigkeit der Potentialvariation auf der Oberfläche von der Dotierdichte in den n-dotierten Bereichen. Für kleine Dichten ergibt sich ein linearer Zusammenhang, für große ist dieser logarithmisch. Die Güte der Anpassung der Modelle (Linien) an die Daten (Symbole) ist nicht perfekt, wird im Weiteren allerdings nicht beachtet.

Die typische mittels Tapping-Mode gemessene Topografie der Probenoberfläche ist in Abbildung 5.17 dargestellt. Augenscheinlich ergibt sich, gut im ebenfalls dargestellten Profil ersichtlich, eine leichte Höhenvariation im Bereich eines Nanometers. Sehr wahrscheinlich sind diese durch Reste der Maske oder durch anisotropes Ätzen während der Maskenentfernung entstanden (die elektrischen Felder in Größenordnung von 10000 V/cm nahe der p-n-Übergänge mögen dafür verantwortlich sein).

Das an den Proben gemessene Oberflächenpotential, beispielhaft in Abbildung 5.18 für $N_D=10^{20} \text{ cm}^{-3}$ dargestellt, zeigt tatsächlich eine Variation mit der Periodizität des erzeugten Gitters. Allerdings sind zwei unerfreuliche Effekte zu beobachten: Zum einen ergeben sich einige Übersprecheffekte zwischen Oberflächenrelief und gemessenem SKM-Potential und zum anderen erscheint die mit etwa 20 mV unter Standardbedingungen ermittelte Potentialvariation deutlich zu gering (für diese Probe wurden etwa 250 mV) erwartet.

Ausgewählte experimentelle Ergebnisse

Um den Einfluss der Oberflächenstruktur auszuschließen, wurden weitere Reste der oxidhaltigen Maske mit Flusssäure geätzt. Im Ergebnis wurden sowohl die Proben als auch die SKM-Karten glatter, eine maßgebliche Änderung der Werte ergab sich allerdings nicht. Ein erster Ansatz zur Erklärung der deutlich zu gering bestimmten Potentialvariation ist die Annahme einer Oberflächenladung. Diese könnte in den p-dotierten Bereichen zu Potentialanstieg und damit zu Abnahme der Differenz führen. Tatsächlich ergeben aber die hier nicht präsentierten Simulationen dazu im üblichen Wertebereich von etwa 10^{10} Ladungen pro cm^2 kaum merkliche Änderung der Potentiale, was in der hohen Dichte der Dotierstoffe begründet ist.

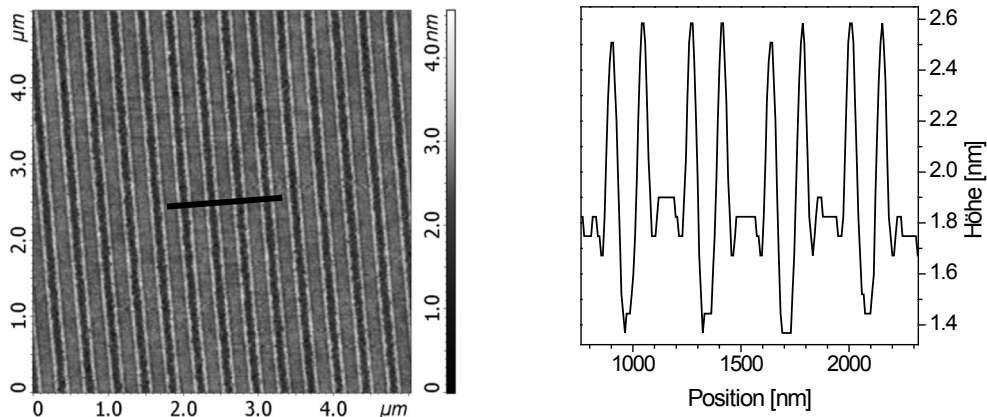


Abbildung 5.17: Topographie der Oberfläche (links) und extrahiertes Profil entlang der angegebenen Linie (rechts). Offenbar sind leichte Höhenvariationen vorhanden, mit großer Wahrscheinlichkeit durch Reste der Maske verursacht.

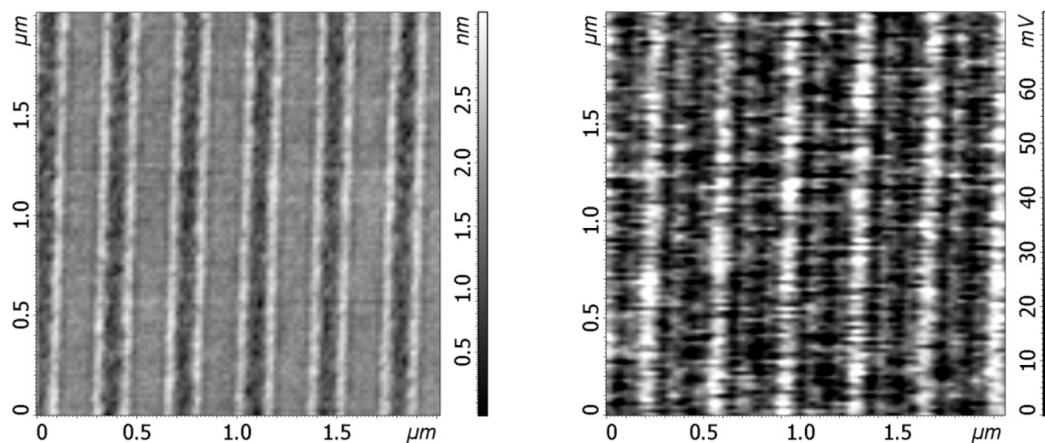


Abbildung 5.18: Topographie (links) und simultan gemessene SKM-Karte einer mit Dotiergitter versehenen Probe. Die Variation des Oberflächenpotentials zeigt einige Artefakte des Reliefs, korreliert aber gut. Allerdings ist der ablesbare mittlere Wert mit etwa 20 mV deutlich unterhalb der erwarteten 250 mV.

Beeinflusst von den theoretischen Vorhersagen bezüglich der Mittelung der Werte wurde als nächstes versucht, durch Variation des Messabstandes zwischen Spitze und Probe im SKM-Durchgang zu extrapolieren. Wie in Kapitel 3 und im vorigen Abschnitt gezeigt, sollte Auftragung der Messwerte über dem Messabstand eine Abschätzung des nicht gefalteten Wertes des Oberflächenpotentials erlauben. Diese Entfaltungsprozedur arbeitet in der Tat sehr gut (Abbildung 5.19, links), allerdings liegend die extrapolierten Werte immer noch etwa 100 mV

unterhalb der erwarteten Werte (Abbildung 5.19, rechts). Der Wert der Differenz variiert dabei mit der Dotierungsdichte, aufgrund der Ungenauigkeiten soll nur der Trend angegeben werden. So ist zu bemerken, dass die Differenz bei kleinen Dichten grösser ist, als bei großen.

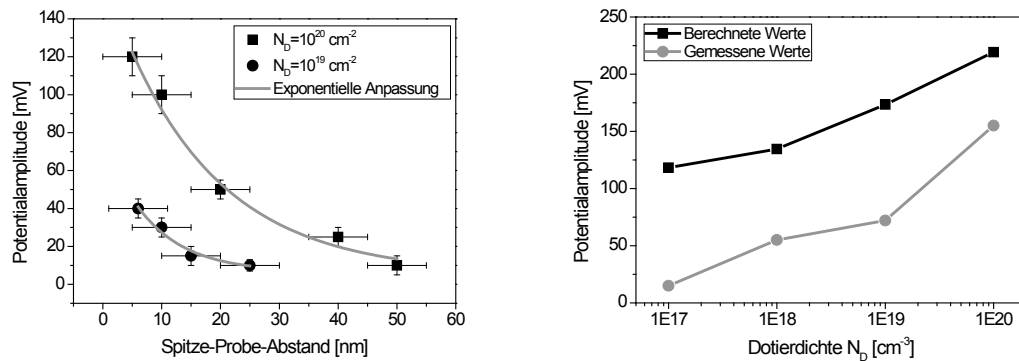


Abbildung 5.19: Darstellung der Messwerte, hier der Potentialvariation, für verschieden stark dotiert Proben in Anhängigkeit der Separation zwischen Sonde und Oberfläche (links). Der exponentielle Zusammenhang erlaubt die Ermittlung extrapolierte Werte, welche den Einfluss der lateralen Mittelung im Experiment unterdrücken. Die erhaltenen Werte sind mit den Vergleichswerten in der rechten Grafik dargestellt.

Diese Erkenntnis, natürlich vor dem Hintergrund der Ergebnisse der gebondeten Wafer, lassen den Schluss zu, dass auch bei diesen Proben die Photogeneration von Ladungsträgern eine Rolle spielt. Die so erzeugte Photospannung könnte maßgeblich den Messwert im SKM-Modus bestimmen.

Um diesen Effekt abzuschätzen, wurden die Drift-Diffusions-Gleichungen für die entsprechenden Strukturen numerisch gelöst. Es stellt sich heraus, dass das Vorgehen, wahrscheinlich begründet durch die hohen lateralen elektrischen Felder, rechnerisch extrem instabil ist. Deshalb wurden die Rechnungen schrittweise durchgeführt, das heißt, zunächst wurde die „Dunkel“-Lösung durch Behandlung der Poissongleichung bestimmt und mit dieser Lösung als Startwert parametrisch die Beleuchtungsstärke erhöht. Einige Ergebnisse dieser Prozedur zeigt Abbildung 5.20.

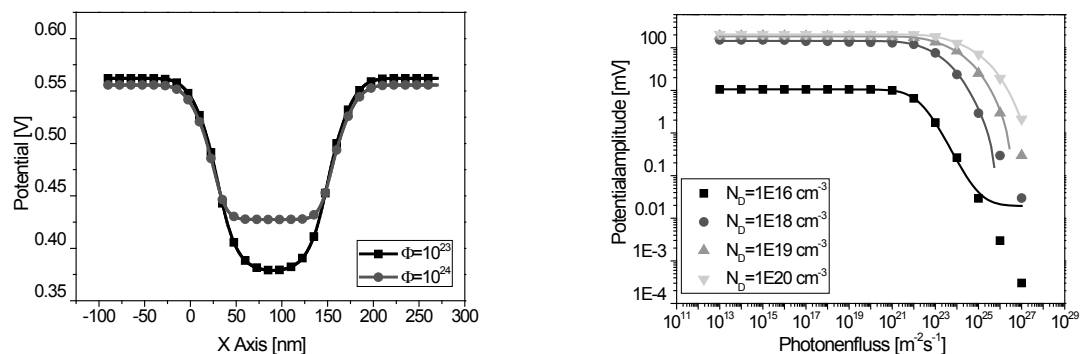


Abbildung 5.20: Numerische Auswertung des Beleuchtungseffekts auf die Dotiergitter. Links ist die Änderung des Potentialverlaufs entlang der Oberfläche beispielhaft für verschiedene Photonenflüsse Φ dargestellt. Der rechten Grafik entnimmt man die berechneten Werte (Symbole) der Potentialdifferenzen für verschiedene Dotierungen in Abhängigkeit der Beleuchtung. Offenbar ergibt sich ein genereller Trend, der auf erzeugte Photospannungen zurückzuführen ist. Die ebenfalls angegebenen Anpassungskurven (Linien) gemäß (C.17) reproduzieren die Daten gut.

Ausgewählte experimentelle Ergebnisse

Wie vermutet, sinkt die Potentialdifferenz stark nichtlinear mit dem Photonenfluss. Anpassungen an entsprechende Photospannungskurven sind in Abbildung 5.20 (rechts) ebenfalls als Linien eingetragen, die Übereinstimmung ist gut. Entnimmt man der Grafik die berechneten Werte für verschiedene Photonenflüsse, lässt sich eine Abbildung der unter Beleuchtung erwarteten Potentialdifferenzen erstellen (Abbildung 5.21). Augenscheinlich gelingt es sehr gut, die experimentell gewonnen Werte numerisch nachzustellen, wenn ein Photonenfluss von etwa $10^{24} \text{ m}^{-2} \text{ s}^{-1}$ angenommen wird, ein Wert also, der tatsächlich im genutzten System vorliegt (die Intensität der Strahlung des Detektionslasers auf der Probenoberfläche liegt etwa bei $I=2 \text{ MW/m}^2$). Aufgrund dieser guten Reproduzierbarkeit ist somit davon auszugehen, dass auch im Fall dieser Proben der Einfluss des Detektionslasers die Messwerte limitiert. Allerdings bleibt darauf hinzuweisen, dass nicht alle Effekte in der Simulation berücksichtigt wurden. So wurde zum Beispiel die Größe der Domäne relativ klein gewählt und Oberflächenrekombination wurde unterdrückt.

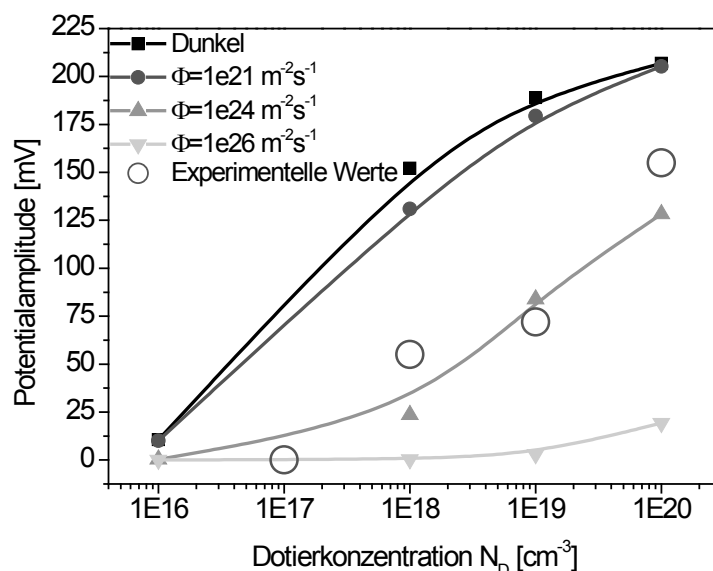


Abbildung 5.21: Verlauf der Potentialamplitude, also der Differenz der Oberflächenpotentiale in n- und p-dotierten Bereichen, in Abhängigkeit der Beleuchtung. Nimmt man einen realen Wert für die Beleuchtungsstärke an, können die experimentell gewonnenen Daten reproduziert werden.

Ebenfalls wurden an diesen Proben Untersuchungen der kapazitiven Signale vorgenommen. Wie Abbildung 5.22 zu entnehmen ist (dargestellt ist die Probe mit $N_D=10^{18} \text{ cm}^{-3}$), ergibt sich ein gut messbarer Kontrast. Einschränkend ist leider zu sagen, dass an den Proben geringerer Dotierung kein topografieunabhängiger Kontrast zu messen war, was wahrscheinlich auf entweder zu geringe Sensitivität des Setups (siehe Abbildung 3.30) oder auf die Maskierung der Ladungsträgerdifferenz durch die Beleuchtung zurückzuführen ist.

Vergleicht man die Korrelationen von Topografie und Potential beziehungsweise Topografie und kapazitivem Signal erkennt man, dass die Bereiche höheren Potentials ebenfalls ein höheres Signal im SCM-Modus erzeugen. Das wird aus den grundlegenden Betrachtungen sofort offensichtlich, da die hoch n-dotierten Bereiche unter den gemachten Voraussetzungen flachere Kapazitäts-Spannungs-Kurven erzeugen.

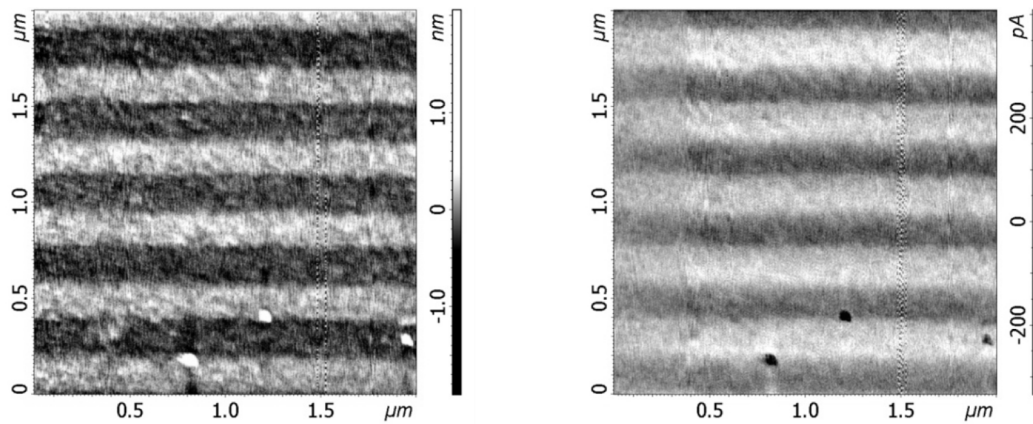


Abbildung 5.22: Topografie (links) und simultan bestimmte SCM-Karte (rechts) eines Dotierungsgitters in Silizium. Offensichtlich erzeugen die, durch SKM identifizierten, n-dotierten Bereiche ein höheres Signal.

Interessant ist die Separationsabhängigkeit des kapazitiven Signals. Wie in Abbildung 5.23 (links) dargestellt ergibt sich eine deutliche nichtlineare Abhängigkeit des Kontrastes, hier der Differenz der Werte in p- und n-dotierten Bereichen, von der Separation. Wie in Kapitel 3 gezeigt, können zwei Prozesse zur Dämpfung von periodischen Variationen in der gemessenen Probenkapazität führen: (a) laterale Mittelung und (b) Zunahme des Verhältnisses zwischen Oberflächen- und Sondenkapazität. Die Signaturen dieser Prozesse sind unterschiedlich, man würde exponentielle Abhängigkeit (a) beziehungsweise Abhängigkeit in Form einer Potenzfunktion (b) erwarten. Wie die in der Abbildung eingefügten Linien illustrieren, ist es möglich, den Signalverlauf mittels beider Prozesse zu beschreiben. Eine Entscheidung darüber zu finden, welcher Effekt die Messung dominiert, ist leider schwierig.

Ebenfalls in Abbildung 5.23 (rechts) dargestellt ist die Spannungsabhängigkeit des Signals. Die Form der Kurve lässt dabei auf Abbildung der kapazitiven Eigenschaften der p-dotierten Bereiche vermuten, da sie den typischen Antwortfunktionen aus Kapitel 3.4 gleicht. Dieser Effekt ist relativ verblüffend, da der Kontrast, definiert als Differenz der Signale aus n- und p-Bereichen doch den entgegengesetzten Verlauf zeigen sollte.

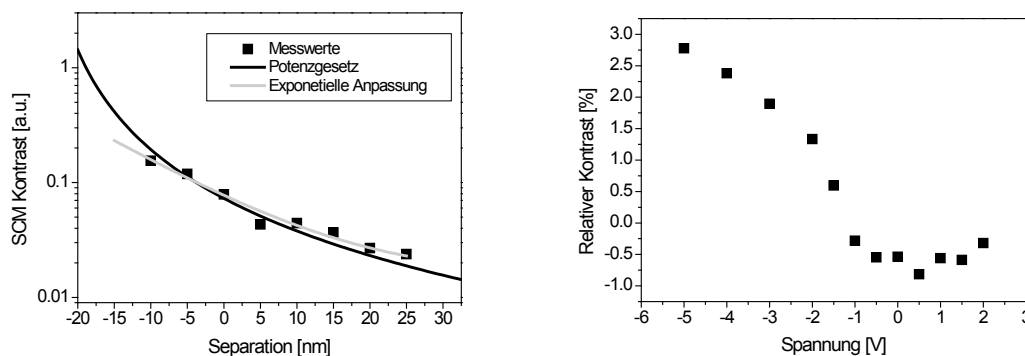


Abbildung 5.23 (links) Abhängigkeit des kapazitiven Kontrastes gemessen an Dotierungsgittern in Silizium von der Separation. Es kann leider nicht entschieden werden, ob laterale Mittelung oder Änderung des Kapazitätsverhältnisses zwischen Sonde und Probe den Verlauf dominieren. (rechts) Abhängigkeit des Kontrastes für feste Separation von der an der Sonde angelegten Spannung. Die Form ähnelt vom Verlauf her den typischen Antwortkurven für die Methode (Kapitel 3.4) spiegelt somit wahrscheinlich die spannungsabhängige Kapazität der p-dotierten Bereiche.

Ausgewählte experimentelle Ergebnisse

Da offenbar aufgrund der kleinen Abmaße der Strukturen eine eindimensionale Beschreibung nicht adäquat ist, wurde die Messungen in dreidimensionaler Simulation nachgestellt. Dazu wurde, um die Komplexität zu verringern, lediglich der Apex der Sonde betrachtet und die n-dotierten Streifen wurden zu einer durch einen p-Bereich unterbrochenen Schicht zusammengefasst. So erhaltene Karten für feste Sondenparameter (hier $R_{Apex}=50$ nm, $S=25$ nm) und unterschiedliche Dotierung des n-Bereiches (links $N_D=10^{18}$ cm⁻³, rechts $N_D=10^{20}$ cm⁻³) zeigt Abbildung 5.24.

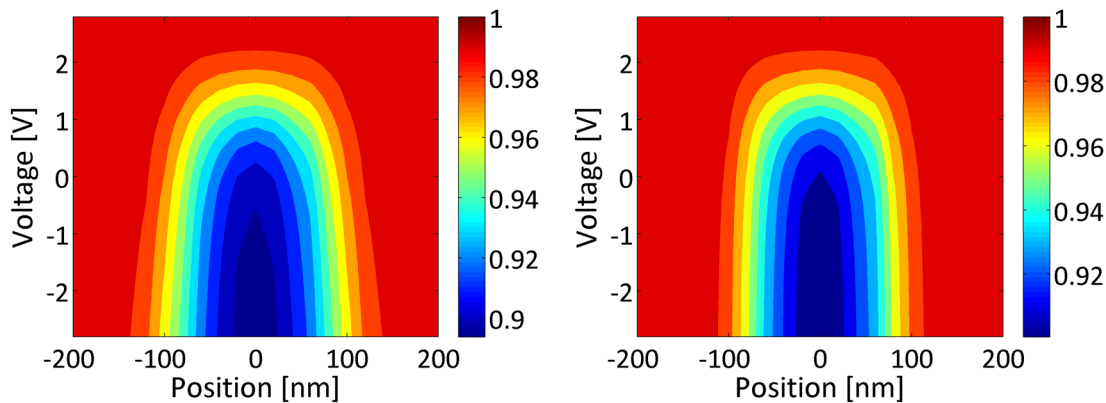


Abbildung 5.24: Numerisch bestimmte spektroskopische SCM-Karten. Die in der Mitte befindlichen p-dotierten Bereiche erzeugen wie erwartet niedrigere Signale. Der Einfluss der Dotierdichte im n-Bereich (links 10^{18} cm⁻³, rechts 10^{20} cm⁻³) ist dabei relativ gering.

Der Abbildung ist zu entnehmen, dass der messbare Kontrast während eines Scans, also horizontale Linien in der Grafik, wenig von der Dotierung beeinflusst werden. So ergibt sich bei beiden angesetzten Dichten etwa derselbe maximale Kontrast von 10%. Lediglich der Verlauf der virtuellen Scans ist etwas anders, was sich durch die unterschiedlichen DebyeLängen erklären lässt. Interessanter ist allerdings der Fakt, dass der p-dotierte Bereich in der Mitte der Darstellungen bei positiven Spannungen ein höheres kapazitives Signal liefert. Somit hat sich durch die Invertierung, welche durch die Übergänge erzeugt wird, auch die Antwortkurve vom p-Typ zum n-Typ geändert. Es werden also damit die Eigenschaften eines n-dotierten Bereichs gemessen, welcher die gleiche Dotierungsdichte aufweist.

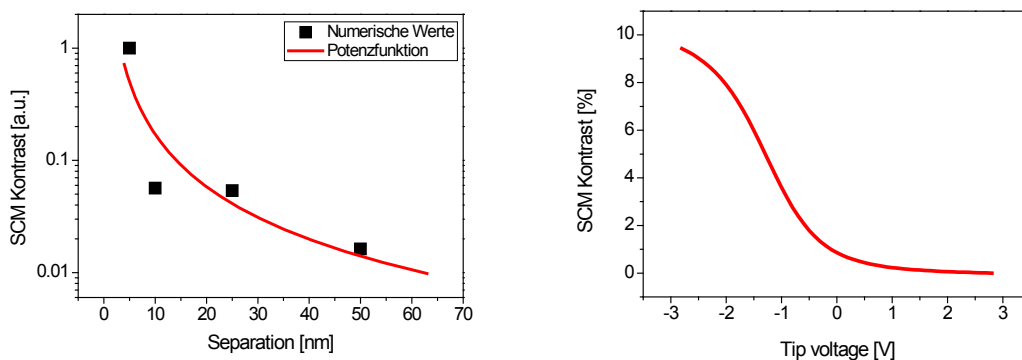


Abbildung 5.25: Aus den Simulationen extrahierte Abhängigkeiten des SCM-Kontrastes für Dotiergitter. Die berechneten Kurven ähneln den experimentellen stark (vgl. Abbildung 5.23), leichte Abweichungen ergeben sich im Exponenten der Dämpfung und im Wert der spannungsabhängigen Signalvariation.

Extrahiert man aus den Ergebnissen der Simulationen die Signalabhängigkeiten bezüglich Separation und Spannung, ergeben sich die Darstellungen in Abbildung 5.25. Die Übereinstimmung mit den gemessenen Werten (Abbildung 5.23) ist zufriedenstellend, was die Verknüpfbarkeit der Ergebnisse unterstreicht. Allerdings zeigen sich auch die Grenzen der vereinfachten Berechnung, da die Exponenten der Abklingfunktion in Experiment und Nachstellung stark unterschiedlich sind. Daraus folgt für eine quantitative Messung, dass Eichkurven für jedes System bestimmt werden müssen. Dieses Vorgehen ist für die Standard-SCM-Methode ebenfalls zwingend notwendig, stellt somit keinen größeren methodischen Nachteil dar.

Einen sehr viel deutlicheren Kontrast ergibt die Anwendung der 3ω -SCM-Methode, wie in Abbildung 5.26 illustriert. Im Gegensatz zu den bisher diskutierten Ergebnissen zeigen sich hierbei die p-dotierten Bereiche stärker. Die Kontrastverstärkung gegenüber der SCM-Methode beträgt hier beinahe 1000%, eine Verstärkung also um einen Faktor 10. Dieses Verhalten wird bei Betrachtung der eindimensional abgeschätzten Werte aus Abbildung 3.41 klar: Während das kapazitive Signal nur um 10% variiert, falls von Dotierdichten in der gegebenen Größenordnung ausgegangen wird, ändert sich das zur Ableitung der Kapazität proportionale um quasi 100%. Die Kontrastumkehr verglichen mit der vorher beschriebenen Methode ergibt sich ebenfalls aus genannter Darstellung.

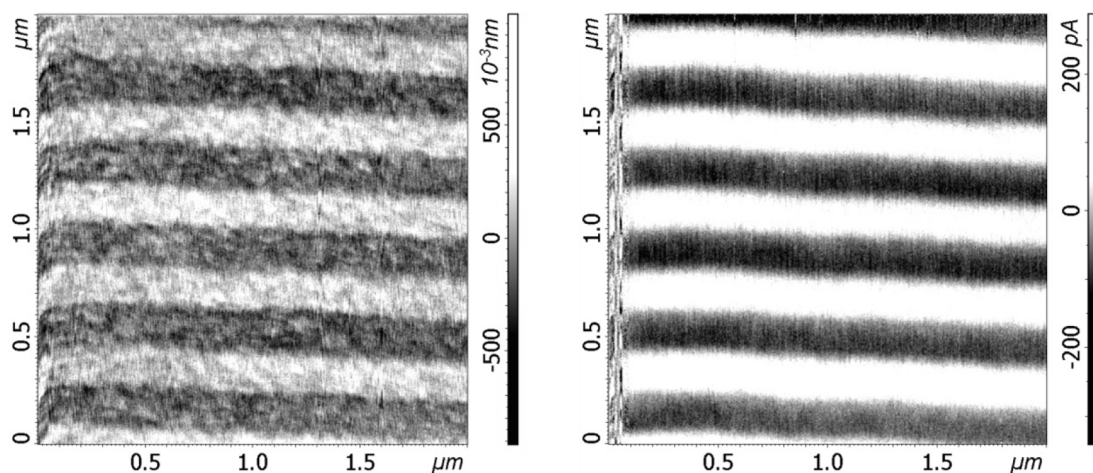


Abbildung 5.26: Topografie (links) und simultan bestimmte 3ω -SCM-Karte (rechts) eines Dotierungsgitters in Silizium. Im Gegensatz zu den bereits präsentierten Ergebnissen der SCM-Methode erzeugen hier die p-dotierten Bereiche ein höheres Signal.

Die Separationsabhängigkeit des Messsignals ist leider experimentell nicht zugänglich, die Signale verhalten sich instabil. Allerdings gelang es, bei festem Spitze-Probe-Abstand die Spannungsabhängigkeit zu bestimmen, welche in Abbildung 5.27a illustriert ist. Die extrahierten Werte entsprechen vom Betrag her den Erwartungen, hingegen ist es nicht möglich, die Kontrastumkehr zu erklären. Würde man nämlich davon ausgehen, lediglich die Differenzen der entsprechenden großflächigen Antwortfunktionen zu messen (Abbildung 3.31), dürfte dieser Effekt nicht auftreten.

Um Klarheit zu schaffen, wurde auch die Entwicklung dieses Signals numerisch nachgestellt. Die entsprechenden Karten für unterschiedliche Dotierung der n-Bereiche zeigt Abbildung 5.28. Leicht ersichtlich ist dabei, dass der Kontrast tatsächlich den aus den Experimenten gefundenen

Ausgewählte experimentelle Ergebnisse

Wert von nahezu 100% annimmt, wobei das Signal in den p-dotierten Bereichen immer höher erscheint als in den n-dotierten Bereichen. Auch in dieser Simulation zeigt sich, dass sich Inversion eingestellt hat: Das niedrig dotierte p-Material verhält sich vom Verlauf der Signal-Spannungs-Kurve her wie ein n-dotiertes Material ähnlicher Dichte (das maximale Signal stellt sich bei positiven Spannungen ein, nicht bei negativen).

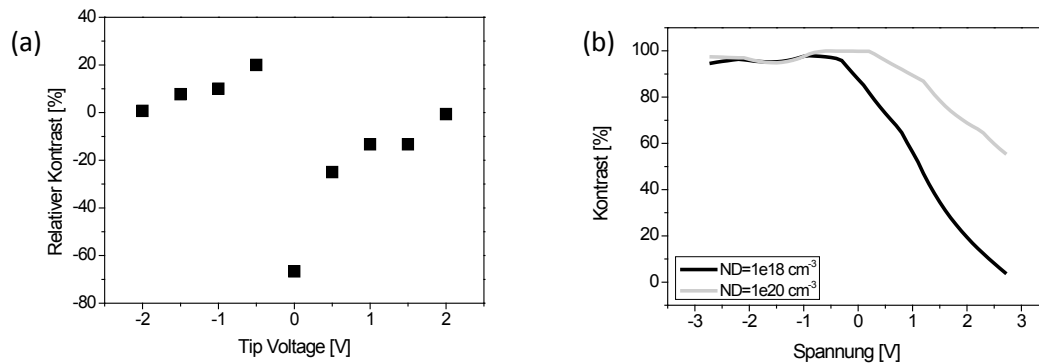


Abbildung 5.27: (a) Spannungsabhängiger Kontrast im 3ω -SCM-Modus. Es ergeben sich Werte im Bereich einiger 10%, allerdings mit Kontrastumkehr. (b) Numerisch bestimmter Kontrast für unterschiedliche Dotierungen: Es lässt sich keine Übereinstimmung der Verläufe feststellen.

Die aus den Karten bestimmten Signal-Spannungs-Kurven sind in Abbildung 5.27b gezeigt. Es ergeben sich zwar die hohen experimentell gefundenen Werte aber auch die numerischen Ergebnisse erklären nicht die Kontrastumkehr. An dieser Stelle muss davon ausgegangen werden, dass mehrere Effekte Anteil an gemessenen Signalen in diesem Modus haben:

1. Offenbar gibt es den erwarteten Dotierungskontrast, der erklärte Mechanismus erscheint richtig.
2. Der Einfluss hoher Wechsellspannungsamplituden, die für die Messung notwendig waren, auf das elektronische System sollte aufgrund der glatten Antwortkurven gering sein.
3. Der Einfluss der Spannungen auf die Cantileverdynamik ist schwer abzuschätzen, da sich sowohl statische Verbiegung als auch Resonanzfrequenz während der Messung ändern können. Möglicherweise überlagern diese Effekte den erwünschten unter Punkt 1 genannten, was letztendlich die quantitative Messung erschwert.
4. Ladungsträgergeneration wurde in der Modellierung unterdrückt. Die möglichen Einflüsse auf die Messsignale sind ebenfalls schwer abschätzbar.

Ein Vergleich der normierten Profile von SKM, 2ω -SCM und 3ω -SCM (Abbildung 5.29) verdeutlicht die Korrelierbarkeit der Signale. Ebenfalls aus dieser Abbildung ersichtlich ist eine laterale Auflösung, die für alle Methoden in der Größenordnung einiger 10 nm liegt. Augenscheinlich unterdrückt zudem die 3ω -Methode einige Artefakte, das Profil erscheint glatter, was geringeren Einfluss der Oberflächenrauigkeit annehmen lässt.

Zusammenfassend kann gesagt werden, dass die in der Arbeit theoretisch vorgestellten Methoden auch experimentell einsetzbar sind, wenn auch mit Abstrichen bezüglich der quantitativen Auswertung. So ist zum Beispiel festzustellen, dass SKM auch bei halbleitenden Proben den Verlauf des Oberflächenpotentials bestimmt, welches hierbei durchaus von

Beleuchtungen der Probe abhängen kann. Die exponentielle Dämpfung periodischer Signale wurde nachgewiesen und die vorgeschlagene Entfaltung erfolgreich angewendet.

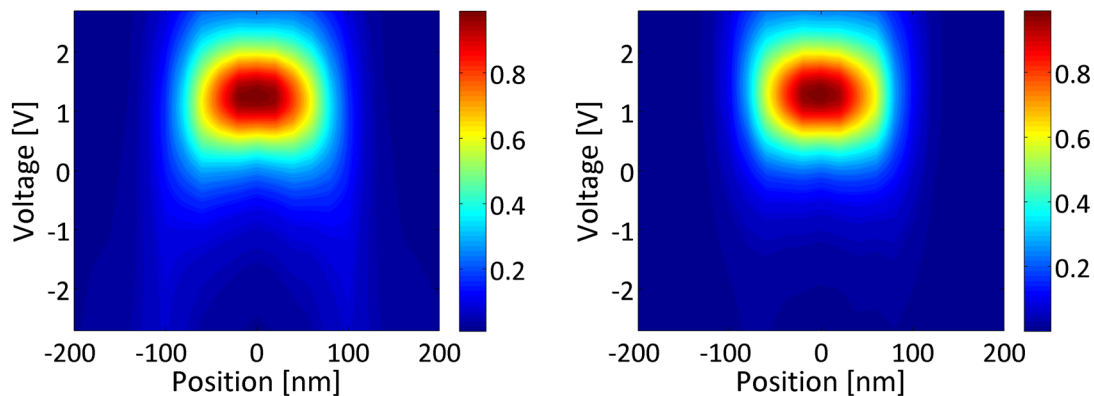


Abbildung 5.28: Numerisch bestimmte spektroskopische 3ω -SCM-Karten. Die in der Mitte befindlichen p-dotierten Bereiche erzeugen höhere Signale. Der Einfluss der Dotierdichte im n-Bereich (links 10^{18} cm^{-3} , rechts 10^{20} cm^{-3}) ist auch in diesem Beispiel gering.

Für die kapazitiven Methoden ergeben sich qualitativ die vorhergesagten Ergebnisse bezüglich Kontrastbildung und Separationsabhängigkeit. Eine zufriedenstellende Erklärung der Spannungsabhängigkeit der Signale ist aber nicht möglich, was generelle Aussagen zur Benutzbarkeit unmöglich macht. Die Erweiterung der Messmethodik auf Abbildung höherer Harmonischer der Anregung erscheint sinnvoll, da zusätzliche Informationen erhalten werden können.

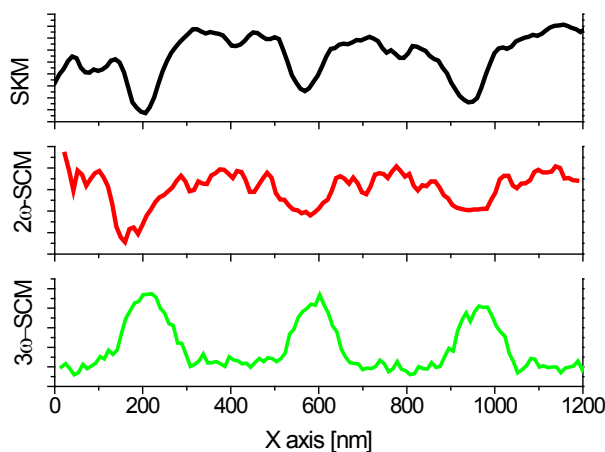


Abbildung 5.29: Vergleich der Signalprofile verschiedener Methoden gemessen an Dotiergittern. Alle Methoden ergeben gut messbaren Kontrast, der Wert dieses Kontrastes und der Einfluss des Oberflächenreliefs variiert allerdings.

Zusammenfassung und Ausblick

Die wichtigsten in der Motivation formulierten Fragestellungen betrafen das Auflösungsverhalten der kraftbasierten Sondenmethoden und deren Informationsgehalt. Diese konnten erfolgreich adressiert werden. So wurde anhand sowohl analytischer Modellrechnungen als auch numerischer Simulation für die Auflösung Folgendes gezeigt:

- a) Das gemessene Signal und die Verteilung eines Potentials auf der Probenoberfläche sind durch eine Faltung verknüpft. Die Faltungsfunktion ähnelt dabei stark der „Poisson-Kernel“-Funktion für die entsprechende Geometrie. Der bestimmende Parameter für die Weite dieser Funktion, und damit für die Auflösung des Experiments, ist die Separation zwischen Sonde und Probe.
- b) Diese Weite hängt wenig von den tatsächlichen Sondenparametern, wie Apexradius und Öffnungswinkel, ab. Vielmehr ist sie für kleine Separationen proportional zur Wurzel der Separation und für größere direkt proportional. Der Übergangspunkt zwischen den genannten Bereichen ist der Radius des Apex.
- c) Aufgrund der funktionalen Form der Transferfunktion lassen sich quasi-eindimensionale Verteilungen experimentell leicht entfalten, da alle vorhandenen Spektralkomponenten der Potentialverteilung exponentiell gedämpft werden.
- d) Prinzipiell gelten die Aussagen auch für kapazitive Signale. Allerdings wird hierbei die exponentielle Dämpfung, welche durch Mittelung entsteht, durch einen stark separationsabhängigen Term überlagert, welcher durch Serienschaltung von Proben- und Sondenkapazität entsteht.
- e) Die zuvor genannten Ergebnisse gelten auch für halbleitende Proben, allerdings mit der Einschränkung reduzierter Ansätze für die Simulation.

Der Informationsgehalt ist, wiederum numerisch bestimmt und analytisch bestätigt, folgender:

- f) Die SKM-Methode misst den Wert des (gemittelten) Oberflächenpotentials unter der Sonde. Dieses Potential kann beeinflusst sein von fixierten Ladungen und, im Fall halbleitender Proben, durch unterschiedliche Dotierung sowie Ladungsträgerdichteveränderungen aufgrund Generation und Rekombination.
- g) Die nc-SCM-Methode erlaubt die Abbildung kapazitiver Eigenschaften, was vor allem bei vorhandenen Schichtstapeln und Dotierungskontrasten wertvolle Informationen liefern kann. Es ist darüber hinaus möglich, den Einfluss der Sondenkapazität zu verringern, wenn Cantileverschwingungen mit höheren Harmonischen der Anregungen zur Messung genutzt werden.

Die vorgestellten experimentellen Ergebnisse unterstreichen diese Annahmen, zeigen allerdings auch die Probleme der Methoden auf:

- h) Obwohl tatsächlich der Potentialverlauf auf der Oberfläche bestimmt werden kann, muss zur Erklärung relativ großer Aufwand betrieben werden. Das liegt darin begründet, dass die Beleuchtung durch den Detektionslaser des AFM-Systems Ladungsträgerdichten generieren kann, die in derselben Größenordnung liegen, wie die der unbeleuchteten Probe. Deshalb sind die entsprechenden Prozesse, wie Oberflächenphotospannung, zu beachten.

Zusammenfassung und Ausblick

- i) Die vorgeschlagene Entfaltungsprozedur greift. Im genutzten zweiphasigen Setup kann demnach die Signalverschmierung effektiv umgangen werden.
- j) Die Interpretation der kapazitiven Signale ist aus den gleichen Gründen schwierig, allerdings zeigt sich zumindest die erwartete Skalierung.
- k) Die quantitative Auswertung der spannungsabhängigen Kontraste gelingt nur teilweise, mögliche Gründe hierfür mögen in der vollständigen Unterdrückung dynamischer Effekte in der Simulation liegen.

Obgleich somit die Zielstellungen erreicht sind, legen die geschilderten experimentellen Probleme folgende zukünftige Anstrengungen nahe:

- l) Modifikation der Geräte, um beleuchtungsabhängige Messungen durchführen zu können. So wäre es wünschenswert, die Gesamtbeleuchtungsstärke variieren zu können, um zum Einen tatsächlich Dunkelwerte messbar zu machen und zum Anderen die Möglichkeit der Abbildung der Oberflächenphotospannung zu methodisieren.
- m) Eine Simulation der Wechselwirkung unter Beachtung der Dynamik wäre wünschenswert. Am jetzigen Punkt erscheint dieser Wunsch nahezu unerfüllbar. Während die Bewegungsgleichungen in Kraftfeldern zumindest bei eindimensionaler Betrachtung numerisch lösbar bleiben, ist die Betrachtung der kompletten beschreibenden Gleichung für Halbleiter extrem aufwendig.

Zusammenfassend bleibt also zu formulieren, dass es im Rahmen dieser Arbeit gelungen ist nachzuweisen, dass kraftbasierte Rastersondenmethoden gut dazu geeignet sind, elektrische Eigenschaften von Proben zu charakterisieren. Wie gezeigt haben die Methoden gute Auflösung und Kontraste, was im Zusammenspiel mit der experimentellen Flexibilität und Einfachheit ihren Einsatz motiviert.

Danksagung

Mein Dank gilt allen Personen in meinem näheren Umfeld, die auf die eine oder andere Art und Weise, zum Beispiel durch Zurverfügungstellung geeigneter Proben und Hilfe bei der Präparation dieser, zur Entstehung dieser Arbeit beigetragen haben.

Dank gilt allen Mitarbeitern unserer Arbeitsgruppe, vor allem Prof. Dr. Reif für die Betreuung und Begutachtung der Arbeit, unserem Techniker Jürgen Bertram für die Unterstützung bei teilweise doch recht schwierigen Bastelaufgaben und Dr. Arguirov für viele hilfreiche Diskussionen.

Großer Dank gebührt den Gutachtern Prof. Dr. Seibold und Dr. Birkholz für ihre Zeit.

Nicht zuletzt danke ich meiner Frau Tanja und meinen Kindern Johanna und Benedikt, die vor allem in der Endphase der Erstellung oft auf mich verzichten mussten und mich dennoch immer moralisch unterstützt haben.

A Grundlegende Gleichungen für die elektrostatische Beschreibung von Silizium

A.1 Die Poissongleichung

Die Poissongleichung lautet in der allgemeinen Form:

$$\Delta\varphi = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \varphi = -\frac{\rho}{\varepsilon} \quad (\text{A.3})$$

Diese Gleichung verknüpft das ortsabhängige elektrostatische Potential φ mit der ortsabhängigen Ladungsdichte ρ . Der Parameter ε ist dabei die Dielektrizitätskonstante des jeweiligen Mediums. Beachtet man, dass die Ladungsträgerdichte aus Dichte der Elektronen, Löcher und der ionisierten Dotieratome besteht, erhält man:

$$\Delta\varphi = -\frac{e(p - n - N_A^- + N_D^+)}{\varepsilon} = -\frac{e(p - n + N)}{\varepsilon} \quad (\text{A.4})$$

Hierbei sind n und p die Dichten von Elektronen und Löchern und N_A^- und N_D^+ die Dichten der ionisierten Dotieratome. Im thermischen Gleichgewicht gilt weiterhin:

$$p = n_i \exp\left(\frac{\varphi_F - \varphi}{V_t}\right); n = n_i \exp\left(\frac{\varphi - \varphi_F}{V_t}\right); V_t = \frac{kT}{q} \quad (\text{A.5})$$

Einsetzen von (A.5) und (A.4) in (A.3), wobei φ_F ein Referenzpotential und V_t die sogenannte thermische Spannung sind, ergibt:

$$\Delta\varphi = \frac{2qn_i}{\varepsilon} \left(\sinh\left(\frac{\varphi - \varphi_F}{V_t}\right) + \sinh\left(\frac{\varphi_F}{V_t}\right) \right) \quad (\text{A.6})$$

φ_F wurde dabei über die Neutralitätsbedingung für den Fall verschwindenden Potentials bestimmt zu:

$$\varphi_F = V_t \sinh^{-1}\left(-\frac{N}{2n_i}\right) \quad (\text{A.7})$$

Für einige Anwendungen ist es günstig, Ort und Potential zu skalieren, um dimensionslose Parameter zu erhalten:

$$u = \frac{\varphi}{V_t}; u_F = \frac{\varphi_F}{V_t}; \bar{X} = \frac{\bar{x}}{\lambda} \quad (\text{A.8})$$

Die Konstante λ lässt sich durch Anwendung der Kettenregel bestimmen:

$$\begin{aligned}
 \frac{\partial}{\partial x} &= \frac{\partial}{\partial X} \frac{\partial X}{\partial x} = \frac{1}{\lambda} \frac{\partial}{\partial X} \\
 \Delta_X \varphi &= V_t \Delta_X u = \frac{V_t}{\lambda^2} \Delta_X u \\
 \frac{V_t}{\lambda^2} \Delta_X u &= \frac{2qn_i}{\varepsilon} (\sinh(u - u_F) + \sinh(u_F)) \\
 \lambda &:= \sqrt{\frac{\varepsilon V_t}{2qn_i}}
 \end{aligned} \tag{A.9}$$

In (A.9) ist Differentiation bezüglich dem tatsächlichen Koordinaten x und den skalierten Koordinaten X durch Benutzung des entsprechenden Index kenntlich gemacht. Die auftretende Skalierungskonstante λ ist die sogenannte intrinsische Debyelänge, welche für Silizium bei $T=300$ K in etwa den Wert $28 \mu\text{m}$ hat.

Für die Poissongleichung in den eingeführten dimensionslosen Variablen erhält man schließlich:

$$\Delta_X u = \sinh(u - u_F) + \sinh(u_F); u_F = \sinh^{-1} \left(-\frac{N}{2n_i} \right) \tag{A.10}$$

A.2 Häufig benutzte Näherungen und typische Längenskalen

Da die Struktur der Poissongleichung hochgradig nichtlinear ist, empfiehlt es sich, natürlich abhängig von der vorgesehenen Anwendung, gewisse Näherungen vorzunehmen. Der Verlauf der im linken Graphen von Abbildung A.1 dargestellten Funktion lässt sich in drei Bereiche einteilen: exponentielles Verhalten für betragsmäßig große Potentiale, lineares Verhalten für kleine Potentiale und fast konstantes Verhalten für mittlere Potentiale.

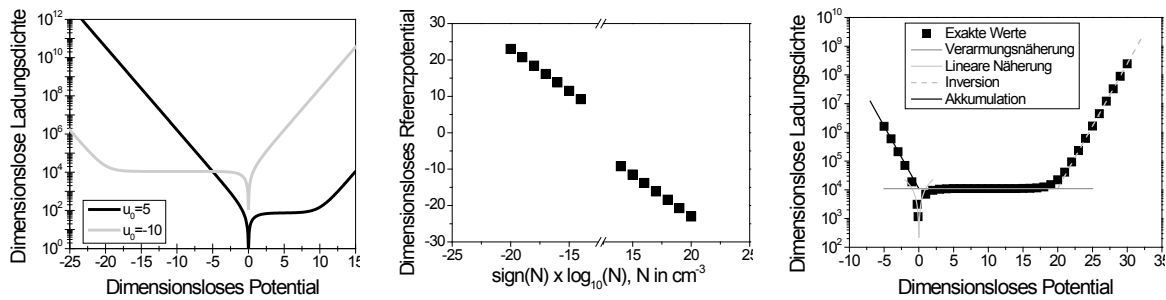


Abbildung A.1: Abhängigkeit der Ladungsdichte (rechte Seite in (A.10)) vom Potential für verschiedene Referenzpotentiale (links), Wert des Referenzpotentials für verschiedene Dotierungsdichten (Mitte) und Verdeutlichung der im Text erläuterten möglichen Näherungen (rechts). Häufig benutzt wird die sogenannte „Verarmungsnäherung“ (rechts, fett gezeichnete Spur), die für moderate Potentiale gültig ist.

Ausgehend von (A.10) können für die genannten Abschnitte folgende Näherungen angegeben werden:

$$\sinh(u - u_F) + \sinh(u_F) \approx \begin{cases} (I): -\frac{1}{2}\exp(u_F - u), u < 0 \\ (II): u \cdot \cosh(u_F), u \approx 0 \\ (III): \sinh(u_F), 1 < u < 2u_F \\ (IV): \frac{1}{2}\exp(u - u_F), u > 2u_F \end{cases} \quad (\text{A.11})$$

„Übersetzt“ in die tatsächlichen Variablen bedeutet das das Folgende: für Spannungen kleiner null ist das Material in Akkumulation, weshalb viele quasifreie Ladungsträger vorhanden sind. Für diesen Fall kann der hyperbolische Sinus durch eine Exponentialfunktion genähert werden und der konstante Term entfällt.

Ist die Spannung maximal um die thermische Spannung V_t von null verschieden, kann die Abhängigkeit der Ladungsträgerdichte von der Spannung als linear angenommen werden, wobei der Proportionalitätsfaktor von der Dotierung abhängt. Mathematisch entspricht dieses Vorgehen einer Taylorentwicklung um den Punkt „0“, wobei nach dem linearen Term abgebrochen wird.

Für den relativ großen Bereich der Spannungen zwischen V_t und dem doppelten der Referenzspannung ist die Ladungsdichte fast konstant gleich der Dichte der Dotierungsatome, man nennt das Material „verarmt“ an freien Ladungsträgern. Um die genannte Näherung zu erhalten, wurde die Ladungsdichte um $u=u_F$ entwickelt, wobei lediglich der konstante Term benutzt wird.

Für größere Spannungen überwiegt wiederum die exponentielle Abhängigkeit von quasifreien Ladungsträgern, das Material ist in Inversion.

Es soll darauf hingewiesen werden, dass eine analytische Lösung des Poissonproblems nur für die Bereiche II und III möglich ist. Sie sollen der Vollständigkeit halber hier angegeben werden:

$$\begin{aligned} (II) \Delta u = \cosh(u_F) \cdot u &\Rightarrow u(X) = u(0) \cdot \exp\left(-\frac{X}{\sqrt{\cosh(u_F)}}\right) \\ (III) \Delta u = \sinh(u_F) &\Rightarrow u(X) = \frac{1}{2}\sinh(u_F) \cdot (X - W)^2, W = \sqrt{\frac{2u(0)}{\sinh(u_F)}} \end{aligned} \quad (\text{A.12})$$

Aus Gleichung (A.12) (II) ergibt sich die Abschirmlänge für kleine Potentiale zu

$\frac{\lambda}{\sqrt{\cosh(u_F)}} \approx \sqrt{\frac{\varepsilon V_t}{q|N|}}$, der sogenannten extrinsischen Debyelänge L_D (beziehungsweise l_D in den dimensionslosen Koordinaten). Die Größe W ist die Ausdehnung der Verarmungszone, welche im Prinzip der intrinsischen Debyelänge entspricht, wenn V_t durch $u(0)$ ersetzt wird.

Obwohl, wie bereits erwähnt, eine vollständige analytische Lösung des Poissonproblems nicht möglich ist, also eine Darstellung des Potentials in Abhängigkeit des Ortes, so kann doch der

Grundlegende Gleichungen für die elektrostatische Beschreibung von Silizium

Zusammenhang zwischen elektrischem Feld und elektrostatischem Potential gewonnen werden. Dazu wird Gleichung (A.10) (in eindimensionaler Schreibweise) mit einem integrierenden Faktor multipliziert und integriert. Man erhält nach Anpassung der Konstanten:

$$\frac{\partial u}{\partial x} = \sqrt{2 \cdot (\cosh(u - u_F) + u \cdot \sinh(u_F) - \cosh(u_F))} \quad (\text{A.13})$$

Das Ergebnis von (A.13) ist nützlich, um zum Beispiel die Flächenladungsdichte an der Oberfläche zu berechnen oder eine weitere charakteristische Längenskala, nämlich die Dicke der Inversionsschicht l_i zu bestimmen:

$$\begin{aligned} u(x) &\sim u(0) \exp\left(-\frac{x}{l_i}\right) \Rightarrow \\ \left. \frac{\partial u}{\partial x} \right|_{x=0} &= \sqrt{2 \cdot (\cosh(u(0) - u_F) + u(0) \cdot \sinh(u_F) - \cosh(u_F))} \approx \frac{u(0)}{l_i} \Rightarrow \\ l_i &\approx \frac{u(0)}{\sqrt{2 \cdot (\cosh(u(0) - u_F) + u(0) \cdot \sinh(u_F) - \cosh(u_F))}} \end{aligned} \quad (\text{A.14})$$

Dabei wurde angesetzt, dass der Verlauf des Potentials exponentiell ist, was für die geringen Abmessungen der Dicke der Inversionsschicht l_i adäquat ist. Die folgende Abbildung A.2 zeigt einen Vergleich der in (A.14) erhaltenen Abhängigkeit mit den Ergebnissen einer numerischen Lösung der Poissongleichung. Dabei wurde für jeden Wert des Potentials an der Oberfläche derjenige Punkt „x“ gesucht, bei dem der angegebene Schwellwert unterschritten wurde. Für den Bereich der Verarmung (hier zwischen Werten von 0 und 20 für das Potential) ist der Verlauf der Kurven identisch; für große und kleine Potential immerhin noch ähnlich: die Schichtdicke nimmt exponentiell mit dem Potential ab. Allerdings sind die auftretenden Abklingkonstanten

stark unterschiedlich: Während Gleichung (A.14) einen Verlauf der Form $l_i \sim e^{-\frac{u}{2.5}}$ erwarten lässt, ergibt sich aus der numerischen Lösung eine höhere Abklingkonstante von etwa 50. Der Ursprung dieser Diskrepanz ist mit großer Wahrscheinlichkeit die Tatsache, dass für (A.14) das maximale elektrische Feld genutzt wurde während die numerische Lösung die Veränderung des Feldes beachtet. Eine Nutzung der Weite der Raumladungszone, definiert als Position an der die Dichte der Ladungsträger auf 1/e des Wertes an der Oberfläche gesunken ist, ergibt eine bessere Übereinstimmung: die Kurven verlaufen parallel, allerdings verschoben um etwa eine Größenordnung. Bei Anwendung eines ähnlichen Vorgehens wie in Gleichung (A.14) lässt sich auch dieses Manko beheben:

$$\begin{aligned} \left. \frac{\partial \rho}{\partial x} \right|_{x=0} &= \frac{\partial}{\partial x} (\sinh(u - u_F) + \cosh(u_F)) = \cosh(u - u_F) \cdot \left. \frac{\partial u}{\partial x} \right|_{x=0} \approx \frac{\rho(0)}{l_i} \Rightarrow \\ l_i &\approx \frac{(\sinh(u(0) - u_F) + \cosh(u_F))}{\cosh(u(0) - u_F) \cdot \sqrt{2 \cdot (\cosh(u(0) - u_F) + u(0) \cdot \sinh(u_F) - \cosh(u_F))}} \end{aligned} \quad (\text{A.15})$$

Die von (A.15) vorhergesagten Werte werden im Bereich der Akkumulation und Inversion sehr gut durch die numerischen Ergebnisse reproduziert.

Abschließend sollen typische Werte für die Referenzpotentiale und Längen angegeben werden:

Tabelle A1: Typische Werte der Referenzpotentiale und Längenskalen für Silizium und $T=300\text{K}$. Die Werte für l_i wurden Gleichung 0.13 entnommen und liegen für die verwendete Spannung etwa eine Größenordnung unter der Debyelänge.

$N_A [\text{cm}^{-3}]$	ϕ_F / u_F	L_D / l_D	$l_i / l_i \text{ für } u=2u_F+5$
10^{15}	285 mV / 11.5	128 nm / 0.0045	16 nm / $6e-4$
10^{17}	405 mV / 16.2	13 nm / 0.0004	1.6 nm / $6e-5$
10^{19}	520 mV / 20.7	1.3 nm / 0.00004	0.16 nm / $6e-6$

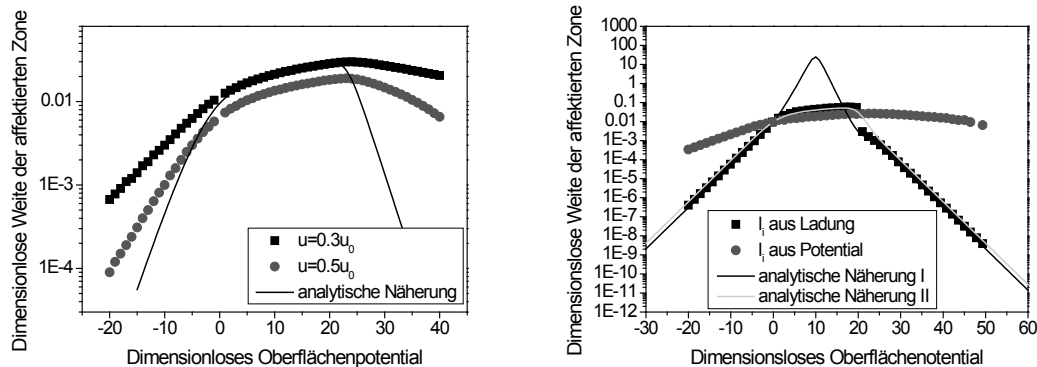


Abbildung A.2: Links: Weite der durch das Oberflächenpotential beeinflussten Zone erhalten durch analytische Näherung (Linie) und numerische Lösung der Poissongleichung und Schwellwertsuche (Symbole). Obwohl der Verlauf der Kurven ähnlich ist, treten große Unterschiede im Bereich der Akkumulation ($u < 0$) und Inversion ($u > 20$) auf, die mit der Einfachheit der Abschätzung (A.14) erklärt werden können. Wird allerdings die Weite der Raumladungszonen als Kriterium genutzt (rechts), ist die Übereinstimmung der Werte viel besser (schwarze Linie). Eine fast komplette Reproduktion erhält man durch Einführung der Oberflächenkapazität (graue Linie, siehe (A.15)).

Wichtig für technische Anwendungen sind üblicherweise nicht das elektrische Feld oder die Weite der beeinflussten Zone, sondern abgeleitete Größen wie zum Beispiel die Kapazität. Eine einfache analytische Konstruktion soll im folgenden Abschnitt aufgezeigt werden:

A.3 Folgerungen aus der Lösung – Die Kapazität der Oberfläche

Die Kapazität einer Halbleiteroberfläche, beziehungsweise eines Metall-Halbleiter-Systems, lässt sich mit Hilfe der bereits gefundenen Formel für das elektrische Feld an der Oberfläche bestimmen:

$$c_s = \left| \frac{\partial}{\partial u_s} \sigma_s \right| \sim \left| \frac{\partial}{\partial u_s} e_s \right| = \left| \frac{\sinh(u_s - u_F) + \sinh(u_F)}{\sqrt{2 \cdot (\cosh(u_s - u_F) + u_s \cdot \sinh(u_F) - \cosh(u_F))}} \right| \quad (\text{A.16})$$

Dabei ist e_s das dimensionslose elektrische Feld an der Oberfläche und u_s das dimensionslose Oberflächenpotential. Nimmt man weiter an, dass die so gewonnen Kapazität der eines Plattenkondensators mit variablem Abstand der Platten entspricht, erhält man für die Dicke dieses Kondensators:

$$c_S = \frac{1}{l_i} \Rightarrow l_i = \left| \frac{\sinh(u_S - u_F) + \sinh(u_F)}{\sqrt{2 \cdot (\cosh(u_S - u_F) + u_S \cdot \sinh(u_F) - \cosh(u_F))}} \right|^{-1} \quad (\text{A.17})$$

Eine grafische Darstellung dieser charakteristischen Länge ist in Abbildung A.2 dargestellt (rote Spur). Besteht das System zusätzlich aus einer dielektrischen Schicht zwischen Halbleiter und Metall (siehe Abbildung 3), muss der Potentialabfall im Dielektrikum beachtet werden:

$$u_G = u_S + t_D \frac{\varepsilon_{r,HL}}{\varepsilon_{r,D}} e_S$$

$$c = \frac{d}{du_G} \sigma_G = \frac{d}{du_G} \sigma_S = \frac{\varepsilon_{r,HL} de_S}{d \left(u_S + t_D \frac{\varepsilon_{r,HL}}{\varepsilon_{r,D}} e_S \right)} = \frac{\varepsilon_{r,HL} \frac{\partial e_S}{\partial u_S} du_S}{du_S + t_D \frac{\varepsilon_{r,HL}}{\varepsilon_{r,D}} \frac{\partial e_S}{\partial u_S} du_S} \quad (\text{A.18})$$

$$c = \frac{\varepsilon_{r,HL} \frac{1}{l}}{1 + t_D \frac{\varepsilon_{r,HL}}{\varepsilon_{r,D}} \frac{1}{l}} = \frac{c_S c_D}{c_D + c_S}$$

In (A.18) bedeuten u_G das Potential am Gate-Kontakt, σ_g und σ_s die Ladungsdichten an Gate und Halbleiteroberfläche und t_D die Dicke des Dielektrikums. Es muss darauf hingewiesen werden, dass das erhaltene Ergebnis sowohl in reellen als auch in normierten Einheiten gültig ist. Allerdings sind die Kapazitäten nicht gemäß (A.17) definiert, sondern beziehen die jeweiligen relativen Dielektrizitätskonstanten mit ein.

Ein direkter analytischer Ausdruck für die Kapazität eines MOS-Systems in der Form $c(u_G)$ ist nicht zugänglich, vielmehr sollte man u_S als Parameter benutzen und daraus u_G und c gemäß (A.18) berechnen. Typische Verläufe von Potential und Kapazitäts-Spannungs-Kurven in realen Systemen zeigt Abbildung A.3. Wichtigste Erkenntnis ist dabei das ausgeprägte Minimum in der Kapazität.

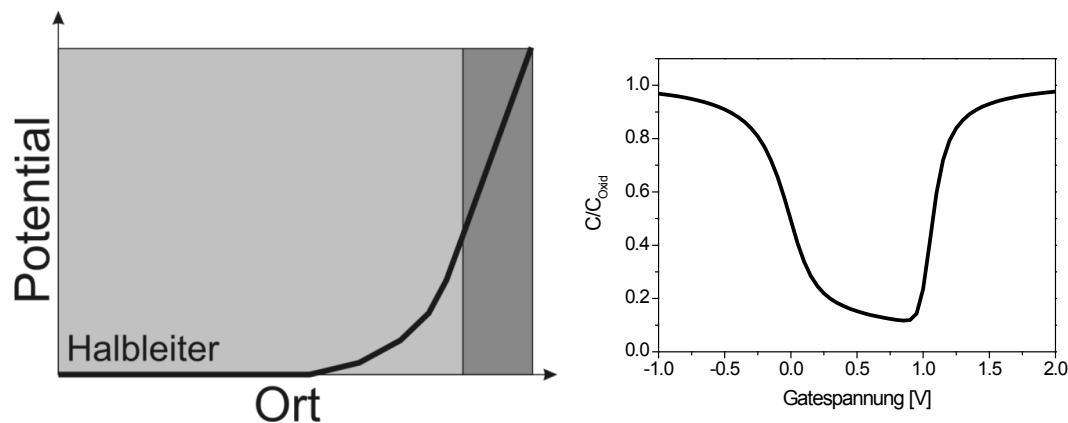


Abbildung A.3: Links: Potentialverlauf in einem MOS-System. Das elektrische Feld steigt im Halbleiter (grau) an während es im Isolator (dunkel) konstant ist. Rechts: Typischer Verlauf der Kapazitäts-Spannungs-Kurve.

A.4 Der Readzylinder

In Kapitel 5, Abschnitt 2 wird über Untersuchungen an gebondeten Proben berichtet. Zur elektrostatischen Beschreibung wird dabei das Modell des Readzylinders herangezogen. Während die Herleitung in der Literatur [38] zu finden ist, sollen hier einige Konsequenzen besprochen werden. Geht man davon aus, das genannte Konstrukt in einer gewissen Tiefe t unter der Oberfläche liegt, kann man durch Überlagerung des Potentials der Linienladung und des Potentials des geladenen Zylinders einen Ausdruck für die Abhängigkeit der Spannung entlang der x-Achse erhalten:

$$V(x, t) = -\frac{(1 - x^2 - t^2)}{2} - \log(\sqrt{x^2 + t^2}) \quad (\text{A.19})$$

In Gleichung (A.19) sind die Längenskalen x und t auf den Radius R_{Read} des Zylinders normiert, das Potential auf einen Wert, der die auftretenden Konstanten vereint. Eine Darstellung des Spannungsverlaufs entnimmt man Abbildung A.4 (links). Offensichtlich ergeben sich Funktionen, die Lorentzfunktionen ähneln. Wichtig für den Vergleich mit der Messung ist die Halbwertsbreite, die leider nicht analytisch zugänglich ist. Eine numerische Auswertung ist in Abbildung A.4 (rechts) dargestellt.

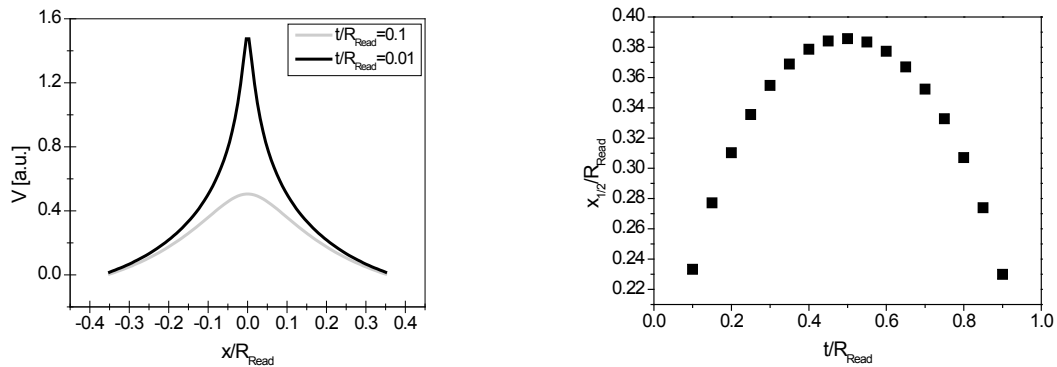


Abbildung A.4: Ergebnisse der Readzylinder-Näherung. Offenbar sind die Potentialverläufe (links) lorentzförmig, die numerisch bestimmten Halbwertsbreiten liegen bei etwa 30% des Radius R_{Read} .

Augenscheinlich liegen die Halbwertsbreiten im Bereich von 30% des Radius des Readzylinders, welcher wiederum im Bereich der Debyelänge des Materials liegt. Zu erwarten wären demnach typische Längenskalen der Potentialverteilungen nahe $1 \mu\text{m}$.

B Intermolekulare Kräfte im System Probe-Sonde

Wie im Kapitel 2 beschrieben, kann die Kraft zwischen zwei Teilchen mit Hilfe des Lennard-Jones-Potentials beschrieben werden. Da allerdings makroskopische Objekte, wie zum Beispiel Sonde und Probenoberfläche aus vielen Teilchen bestehen, muss die Theorie erweitert werden.

B.1 Anziehende Kräfte

Zunächst wird der „anziehende“ Teil des Potentials besprochen, also der Bereich in dem Dipol-Dipol-Wechselwirkungen zum Tragen kommen. Potential und Kraft stellen sich wie folgt dar:

$$\begin{aligned} V &= -2\varepsilon \left(\frac{r_0}{r} \right)^6 \Rightarrow \tilde{V} = -\frac{1}{\tilde{r}^6} \\ F &= -\frac{12\varepsilon}{r_0} \left(\frac{r_0}{r} \right)^7 \Rightarrow \tilde{F} = -\frac{1}{\tilde{r}^7} \end{aligned} \quad (\text{B.1})$$

Die eingeführten normierten Variablen ergeben sich durch Division durch die auftretenden Konstanten.

Soll die Gesamtkraft für ein Zwei-Körper-System errechnet werden, muss über die entsprechenden Volumina integriert werden:

$$\vec{F}_{\text{Gesamt}} = \int_{V_1} \int_{V_2} n_1(\vec{r}_1) \cdot n_2(\vec{r}_2) \cdot \frac{12\varepsilon}{r_0} \left(\frac{r_0}{r(\vec{r}_1, \vec{r}_2)} \right)^7 \vec{e}(\vec{r}_1, \vec{r}_2) dV_1 dV_2 \quad (\text{B.2})$$

Darin sind V_1 und V_2 die zu betrachtenden Volumina der Körper, n_1 und n_2 die entsprechenden Teilchendichten sowie r_1 und r_2 die Ortsvektoren der Volumenelemente. Der Vektor \vec{e} stellt den Einheitsvektor der Verbindungsline zwischen den Volumenelementen an r_1 und r_2 dar. Es soll an dieser Stelle darauf hingewiesen werden, dass das anziehende Potential zwischen zwei Molekülen/Atomen alternativ wie folgt angegeben werden kann:

$$F = -\frac{6A}{r^7} \quad (\text{B.3})$$

In dieser Gleichung ist r wiederum der Abstand und A die sogenannte Hamakerkonstante, welche sich aus den Ionisierungspotentialen und Polarisierbarkeiten der beteiligten Spezies zusammensetzt [45]. Die typische Größenordnung für A ist 10^{-78} Jm^6 .

Einige Sonderfälle des doch recht komplizierten Integrals sollen hier besprochen werden:

1. Punkt – Halbraum

Für dieses stark vereinfachte System (siehe Abbildung B.1) ergibt sich nach Transformation in Zylinderkoordinaten folgendes Integral zur Bestimmung der Gesamtkraft in einer ausgezeichneten Richtung (hier die z-Achse):

Intermolekulare Kräfte im System Probe-Sonde

$$F_{\text{Gesamt}} = - \int_V n_2(\vec{r}_2) \cdot 6A \frac{\cos(\alpha)}{r(\vec{r}_2)^7} dV = - \int_0^{2\pi} \int_0^\infty \int_{-\infty}^0 n \cdot 6A \cdot \rho \cdot \frac{H-z}{\sqrt{((H-z)^2 + \rho^2)^3}} dz d\rho d\varphi \quad (\text{B.4})$$

Wie auch aus Abbildung B.1 ersichtlich, bedeutet H den Abstand zwischen dem betrachteten Punkt und dem Halbraum und ρ, φ, z sind die benutzten Zylinderkoordinaten. Die Integration über das Volumen des zweiten Körpers wurde hierbei natürlich unterdrückt.

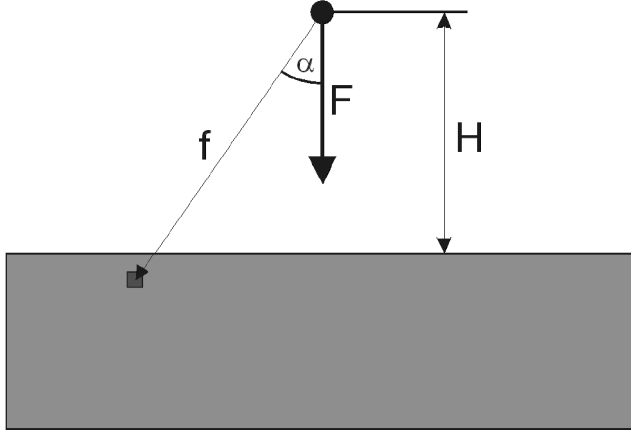


Abbildung B.1: Untersuchte Geometrie im Fall der Wechselwirkung zwischen einem Punkt und einem Halbraum

Die Lösung von 1.4 ergibt sich zu:

$$F_{\text{Gesamt}} = - \frac{\pi \cdot n \cdot A}{2 \cdot H^4} \quad (\text{B.5})$$

2. Halbkugel – Halbraum

Die auf eine Kugel mit Radius R vor einem Halbraum (Abbildung B.2) wirkende Kraft erhält man durch Durchführung des bisher unterdrückten Integrals. Die Rechnung wird vereinfacht, wenn man beachtet, dass alle Punkte auf einer kreisförmigen Schnittebene der Kugel denselben Abstand zur Halbebene haben. Dadurch entfallen zwei der nötigen Integrationsschritte. Die auf eine Kreisscheibe wirkende Kraft ist:

$$f = - \frac{\pi \cdot n^2 \cdot A}{2 \cdot (S+z)^4} \cdot \pi r^2 \quad (\text{B.6})$$

In obiger Gleichung ist S der Abstand zwischen Kugel und Ebene, z die Höhenkoordinate und r der Radius der Kreisscheibe. Des Weiteren wurde der Einfachheit halber angenommen, dass die Teilchendichten gleich sind. Nach Anwendung einfacher geometrische Grundbeziehungen ergibt sich für die Gesamtkraft somit:

$$F_{\text{Gesamt}} = - \int_0^R \frac{\pi \cdot n^2 \cdot A}{2 \cdot (S+z)^4} \cdot \pi \cdot (2 \cdot R \cdot z + z^2) dz = \frac{\pi^2 \cdot n^2 \cdot A}{6} \frac{R^3 (R+4S)}{S^2 (R+S)^3} \quad (\text{B.7})$$

Eine genauere Untersuchung dieses Resultates ergibt, dass für $S < R$ die Kraft invers proportional zum Quadrat von S fällt und für große Abstände $S > R$ die bereits für die Wechselwirkung Punkt-

Ebene gefundene S^{-4} -Abhängigkeit überwiegt. Veranschaulicht wird dieses Verhalten in Abbildung A 7.

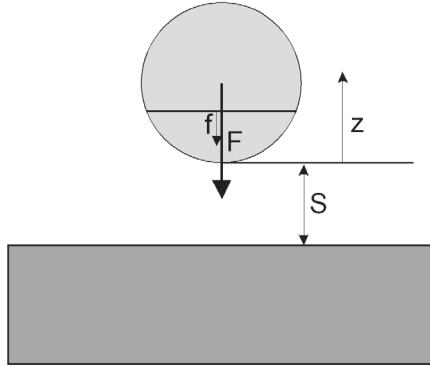


Abbildung B.2: Verwendete Geometrie für die Berechnung der Kraftwirkung auf eine Kugel.

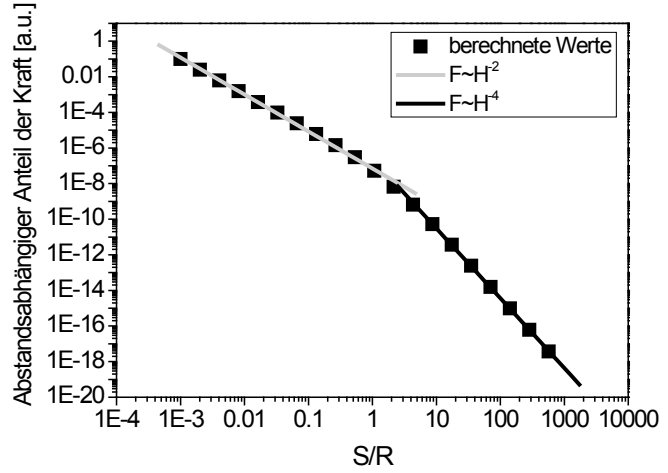


Abbildung B.3: Prinzipielle Abhängigkeit der Kraft auf eine Kugel in Abhängigkeit vom Abstand zur Oberfläche.

3. Konische Sonde mit abgerundeter Spitze

Diese Geometrie kommt den üblicherweise in AFMs benutzten Spitzen sehr nahe und soll deshalb hier besprochen werden. Der Kontaktpunkt zwischen dem kugelförmigen Teil und dem konischen Schaft ist so gewählt, dass der Übergang glatt ist, die ersten Ableitungen der Randkurvenfunktionen gleich sind:

$$z_{Kugel}(r) = S + R - \sqrt{R^2 - r^2}, z_{Schaft}(r) = z_0 + \frac{r - r_0}{\tan(\varphi)} \quad (B.8)$$

$$\frac{dz_{Kugel}(r_0)}{dr} = \frac{r_0}{\sqrt{R^2 - r_0^2}} = \frac{dz_{Schaft}(r_0)}{dr} = \frac{1}{\tan(\varphi)}$$

In Gleichung (B.8) ist φ der Öffnungswinkel des Schafts und r_0 beziehungsweise z_0 die zu bestimmenden Koordinaten des Kontaktpunktes. Es ergibt sich folgende Gleichung für die Kraft:

$$F_{Gesamt} = - \int_0^{z_0} \frac{\pi \cdot n^2 \cdot A}{2 \cdot (S + z)^4} \cdot \pi \cdot (2 \cdot R \cdot z + z^2) dz - \int_{z_0}^H \frac{\pi \cdot n^2 \cdot A}{2 \cdot (S + z)^4} \cdot \pi \cdot (r_0 + \tan(\varphi)(z - z_0))^2 dz$$

$$F_{Gesamt} \approx - \frac{\pi^2 \cdot n^2 \cdot A}{6} \left[\frac{z_0^2 (3SR + z_0R + z_0S)}{S^2 (S + z_0)^3} + \frac{\tan(\varphi)^2 (S + z_0)^2 + r_0 \tan(\varphi)(S + z_0) + r_0^2}{(S + z_0)^3} \right] \quad (B.9)$$

Um die Komplexität zu verringern, wurde das zweite Integral nicht bis zur Höhe H des Schafts, sondern bis ins Unendliche ausgeführt. Wie in Abbildung B.5 ersichtlich, nähert sich die hier errechnete Gesamtkraft für kleine Abstände dem bereits von der Kugel bekannten Verlauf an, während bei größeren Abständen die Kraft invers proportional zum Abstand wird.

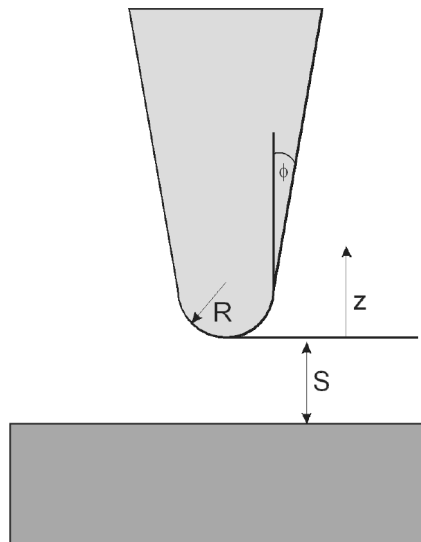


Abbildung B.4: Verwendete Geometrie bei der Berechnung der Kraftwirkung auf eine konische Spitze.

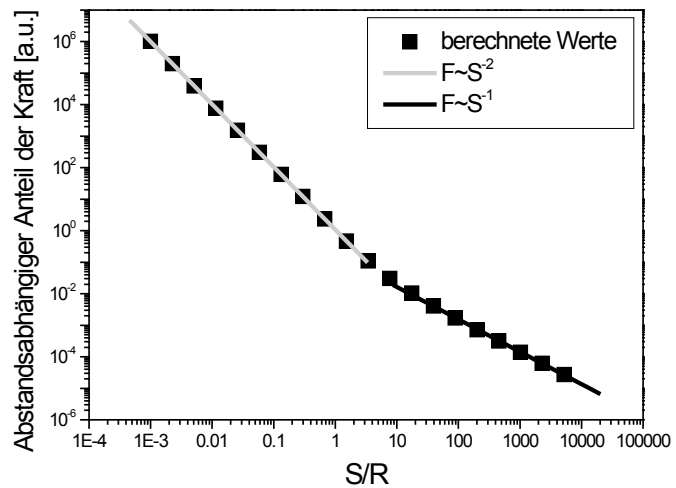


Abbildung B.5: Prinzipieller Verlauf der Kraft auf eine konische Spitze in Abhängigkeit vom Abstand zur Oberfläche.

Abschließend soll die Größenordnung der wirkenden Kräfte abgeschätzt werden. Dazu werden folgende Werte angenommen, welche typischen Geometrien und Materialien angelehnt sind: $n=5 \times 10^{28} \text{ m}^{-3}$, $A=1 \times 10^{-77} \text{ Jm}^6$, $R=30 \text{ nm}$, $\varphi=10^\circ$. Es ergeben sich Kräfte im Bereich einiger nN falls der Abstand im Bereich des Radius ist und Kräfte in der Größenordnung einiger fN für größere Abstände.

Der Vollständigkeit halber soll erwähnt werden, dass die auf den Cantilever wirkende Kraft vernachlässigt werden kann, was folgend Abschätzung zeigt: Da die Dicke d des Cantilevers mit etwa $2 \text{ }\mu\text{m}$ deutlich geringer als der minimale Abstand zur Oberfläche, die Höhe H der Spitze mit etwa $20 \text{ }\mu\text{m}$, ist, kann die wirkende Kraft wie folgt genähert werden:

$$F_{\text{Cantilever}} \approx \frac{\pi r_{\text{Cantilever,effektiv}}^2 d_{\text{Cantilever}}}{(S + H)^4} \quad (\text{B.10})$$

Der hier eingeführte effektive Radius des Cantilevers ergibt sich durch Wandlung des rechteckigen Querschnitts in einen kreisförmigen unter Beachtung der Tatsache, dass eine auf der gesamten Fläche angreifende Kraft bei einem einseitig eingespannten Balken eine geringere Verbiegung als dieselbe Kraft punktförmig am freien Ende wirkend verursacht:

$$r_{\text{Cantilever,effektiv}} = \frac{\sqrt{\frac{3}{8} B \cdot L}}{\pi} \quad (\text{B.11})$$

In obiger Gleichung bedeuten B und L die Breite beziehungsweise Länge des Balkens. Ein Vergleich mit den bereits gefundenen Abstandsabhängigkeiten der Kraft für kleine und größere Abstände (R/S^2 und $\tan^2 \varphi / S$) ergibt, dass die auf den Cantilever wirkende Kraft mindestens eine Größenordnung geringer ist. Da sie zudem nur bei großen Abständen im Bereich der Höhe der Spitze eine rechnerische Rolle spielt und in diesem Bereich die praktische Durchführung eines

AFM-Experiments wenig sinnvoll und auch messtechnisch schwierig erscheint, kann sie vernachlässigt werden.

B.2 Abstoßende Kräfte

Da die Sonde des AFM-Systems während der Messung elektrostatischer Kräfte, wie sie in den dieser Arbeit zugrunde liegenden Experimenten durchgeführt wurden, nicht in Kontakt mit der Probeoberfläche ist, sollen hier lediglich die wichtigsten Grundlagen der mechanischen Kontaktbildung angeführt werden.

1. Das Hertz-Problem

Unter Annahme einiger Einschränkungen wie zum Beispiel Isotropie und geringe Verformung kann die Abhängigkeit der repulsiven Kraft von der Eindringtiefe d wie folgt angegeben werden [46]:

$$F = K d^{\frac{3}{2}} R^{\frac{1}{2}} \quad (\text{B.12})$$

Darin sind K und R effektive elastische Konstanten und Radien der beteiligten Körper, die sich wie folgt berechnen:

$$\begin{aligned} \frac{1}{R} &= \frac{1}{r_1} + \frac{1}{r_2} \\ \frac{1}{K} &= \frac{3}{4} \left(\frac{1 - \mu_1^2}{E_1} + \frac{1 - \mu_2^2}{E_2} \right) \end{aligned} \quad (\text{B.13})$$

In Gleichung (B.13) sind $r_{1,2}$ die Radien der Körper, $\mu_{1,2}$ die Poissonzahlen und $E_{1,2}$ die E-Module der Stoffe. Für den Fall Kugel/Ebene unter Annahme gleicher Stoffe ergeben sich die Größen zu:

$$\begin{aligned} R &= r \\ K &= \frac{2}{3} \frac{E}{1 - \mu^2} \end{aligned} \quad (\text{B.14})$$

2. Das DMT-Modell

Üblicherweise ist es nicht ausreichend, lediglich Phänomene ohne Kontaktbildung oder mit Kontaktbildung zu betrachten, sondern auch Zwischenzustände, bei denen Teile der Körper repulsive Kräfte verspüren während andere attraktiven Kräften ausgesetzt sind. Tatsächlich tritt in diesem Bereich ein neuer Mechanismus auf, die Adhäsion. Diese Kraft wird durch molekulare Kräfte mittlerer Reichweite, wie zum Beispiel Wasserstoffbrückenbindungen oder andere elektrostatische Wechselwirkungen verursacht.

Das DMT-Modell (Derjagin, Muller, Toropov [47]) erweitert das Hertzmodell um die Kraft, die der auftretende Kontaktring aufgrund der Adhäsion verspürt:

$$F = K d^{\frac{3}{2}} R^{\frac{1}{2}} - 2\pi R \sigma \quad (\text{B.15})$$

Der zusätzlich erscheinende Parameter σ steht für die spezifische Adhäsionsarbeit.

Typische Größenordnungen für die auftretenden repulsiven Kräfte sind einige nN bei einem Nanometer Eindringtiefe, die Adhäsionskräfte haben eine vergleichbare Stärke.

B.3 Das komplette Modell

Eine direkte Summation der anziehenden und abstoßenden Kräfte ist nicht möglich, des Weiteren würde eine einfache stückweise Definition zu unstetigen Ableitungen führen. Deshalb ist es zweckmäßig, die Kraft-Abstands-Kurve wie folgt zu erklären:

$$F(S) = \begin{cases} KR^2(a-S)^{\frac{3}{2}} - F_{Min} & S < a \\ p_3(S-a)^3 + p_2(S-a)^2 + p_1(S-a) + p_0 & a \leq S \leq b \\ -\frac{CR}{6S^2} & S > b \end{cases} \quad (B.16)$$

F_{Min} ersetzt dabei die Adhäsionskomponente aus Gleichung(B.15), C fasst die Konstanten aus (B.7) zusammen und p_i sind Koeffizienten des kubischen Polynoms, welches die Kurven stetig differenzierbar in den Punkten a und b verknüpft. Die Lösung der vier Bestimmungsgleichungen für die Koeffizienten ergibt schließlich:

$$\vec{p} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} -F_{Min} \\ 0 \\ -\frac{(5b-2a)CR}{6b^3(b-a)^2} + \frac{3F_{Min}}{(b-a)^2} \\ \frac{(2b-a)CR}{3b^3(b-a)^3} - \frac{2F_{Min}}{(b-a)^3} \end{pmatrix} \quad (B.17)$$

Eine grafische Darstellung des Verlaufs der Kraft-Abstands-Kurve zeigt folgende Abbildung:

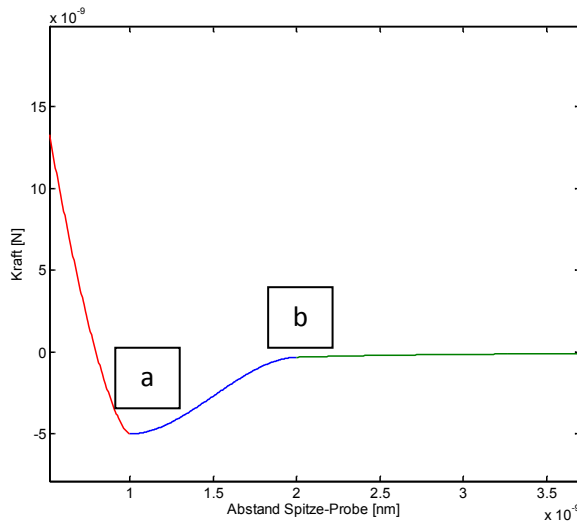


Abbildung B.6: Modellierter Kraft-Abstands-Kurve für typische Parameter. Die Verknüpfungspunkte a und b sind gekennzeichnet.

C Der belichtete Halbleiter

Da im Allgemeinen AFM-basierte Messungen nicht in Dunkelheit durchgeführt werden können, da das System zur Detektion der Cantilverbiegung einen Laser benutzt, erscheint es notwendig, lichtinduzierte Prozesse in Halbleitern zu besprechen. Zu diesen Prozessen gehören Generation und Rekombination von Ladungsträgern sowie ihre Umverteilung mittels Diffusion und feldinduzierte Drift. Des Weiteren ergeben sich, für die tatsächliche Messung ungleich wichtigere, sekundäre Effekte wie die Ausbildung einer Oberflächenspannung.

C.1 Generation und Rekombination

Zur Beschreibung der Photogeneration soll Folgendes angenommen werden:

1. Es gilt das Lambert-Beersche Gesetz.
2. Jedes absorbierte Photon erzeugt ein Elektron-Loch-Paar.

Mathematisch lässt sich die erste Bedingungen wie folgt formulieren:

$$\frac{dI}{dx} = -\alpha I \quad (\text{C.1})$$

Dabei ist I die Intensität, α die (ortsabhängige) Absorptionskonstante und x die Ortskoordinate entlang der Ausbreitungsrichtung. Die Lösung dieser Differentialgleichung ergibt sich zu:

$$I(x) = I_0 e^{-\int_0^x \alpha(z) dz} \approx I_0 e^{-\alpha x} \quad (\text{C.2})$$

In Gleichung (C.2) sind sowohl die Lösung für veränderliche und nichtveränderliche Absorptionskonstante angegeben. Ausgehend von der deponierten Energiedichte (Gleichung (C.1)) unter Berücksichtigung der Energie-Frequenz-Beziehung ergibt sich für die zweite Bedingung:

$$G(x) = \frac{1}{\hbar \omega} \frac{dI}{dx} \approx \frac{\alpha I_0}{\hbar \omega} e^{-\alpha x} = G_0 e^{-\alpha x} \quad (\text{C.3})$$

Hier stehen G für die Generationsrate, also für die Anzahl generierter Ladungsträger pro Volumen und Zeit, ω für die Kreisfrequenz des verwendeten Lichts und \hbar ist das Plancksche Wirkungsquantum. Die Generationstiefe ist somit lediglich von den optischen Parametern des verwendeten Materials abhängig, als Beispiel sind entsprechende Werte für Silizium in Abbildung C.1 illustriert. Dabei ist neben der Absorptionskonstante auch deren Kehrwert, die Eindringtiefe, angegeben (Werte aus [48]). Der Vollständigkeit halber wird auch der Verlauf der Reflektivität angegeben, da natürlich nur der eindringende Anteil des Lichts einen Einfluss auf die Ladungsträgerdichten haben kann (aus [49]).

Der belichtete Halbleiter

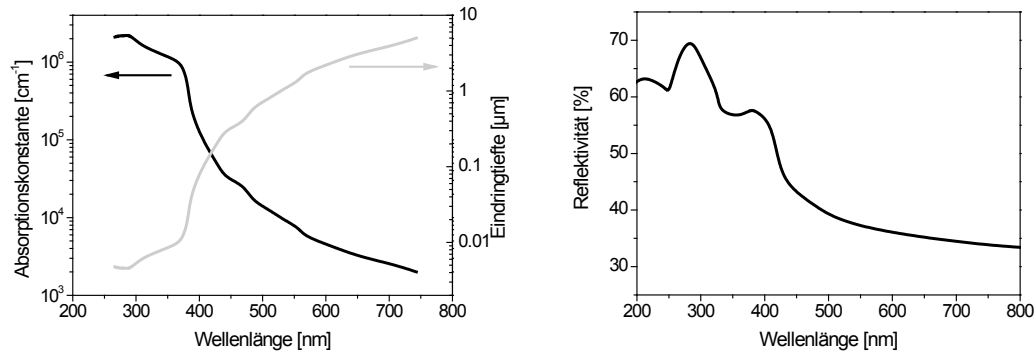


Abbildung C.1: Abhängigkeit der Absorptionskonstanten sowie der sich ergebenden Eindringtiefe des Lichts für Silizium bei Raumtemperatur (linker Graph) und Abhängigkeit der Reflektivität der Oberfläche von der verwendeten Wellenlänge (rechter Graph).

Da die zusätzlich vorhandenen Elektron-Loch-Paare aus energetischen Betrachtungen ungünstig sind, werden verschiedene Prozesse für das Verschwinden dieser verantwortlich sein. Üblicherweise unterscheidet man in: Band-Band-Rekombination (BB), Shockley-Read-Hall-Rekombination (SRH), Auger-Rekombination und Oberflächenrekombination. Schematisch sind diese Prozesse in Abbildung C.2 dargestellt.

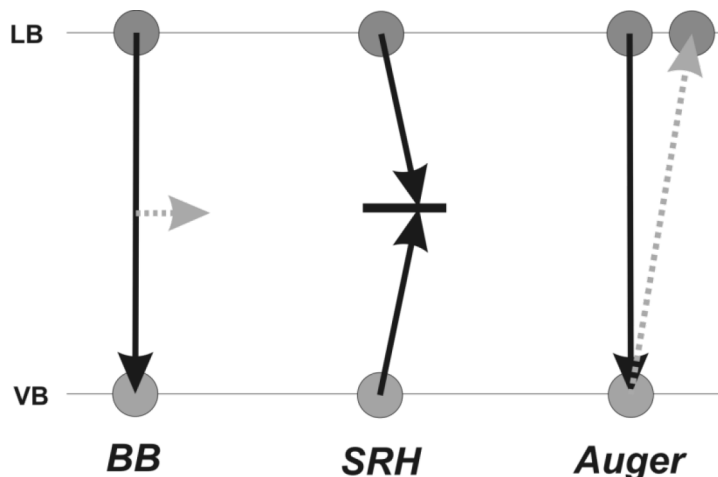


Abbildung C.2: Typische Rekombinationspfade in Halbleitern: Bei der Band-Band-Rekombination (BB) wird die Bandlücke zwischen Valenz- und Leitungsband (VB und LB) komplett überwunden und die Energie als Photon emittiert, bei der strahlungslosen Shockley-Read-Hall-Rekombination (SRH) dient ein lokalisierter Defektzustand als Zentrum und bei der Auger-Rekombination wird die bei einem BB-Übergang freiwerdende Energie an ein drittes Teilchen übergeben.

Die zugehörigen Rekombinationsraten (Anzahl der rekombinierenden Ladungsträger pro Volumen und Zeit) sollen hier lediglich angegeben werden, ausführliche Erläuterungen finden sich in [50]:

1. Band-Band-Rekombination:

$$R_{BB} = b \cdot (np - n_i^2) \quad (\text{C.4})$$

Hierbei ist b der bimolekulare Rekombinationskoeffizient, der für Silizium bei Raumtemperatur etwa $5 \times 10^{-15} \text{ cm}^3/\text{s}$ beträgt.

2. SRH-Rekombination

$$R_{SRH} = \frac{n \cdot p - n_i^2}{\tau_e \cdot (n + n_1) + \tau_p \cdot (p + p_1)} \quad (\text{C.5})$$

$$n_1 = N_C \cdot e^{\frac{E_T - E_C}{kT}}, p_1 = N_V \cdot e^{\frac{E_V - E_T}{kT}}$$

Die hier angegebenen Lebensdauern τ_n und τ_p ergeben sich aus der Charakteristik des Defekts. Da die Oberflächenrekombination prinzipiell vergleichbar abläuft, ist es zweckmäßig, diese beiden Prozesse zu kombinieren. Es ergibt sich eine effektive Lebensdauer zu:

$$\tau_{eff} = \frac{1}{\sigma \cdot N \cdot v + \frac{s}{d}} \quad (\text{C.6})$$

Das Produkt aus Wirkungsquerschnitt σ , Defektdicht N und thermischer Geschwindigkeit v stellt den Kehrwert der Volumenlebensdauer dar, s gibt die Oberflächenrekombinationsgeschwindigkeit an und d ist die Dicke der Probe.

3. Auger-Rekombination

$$R_{Auger} = (c_n \cdot n + c_p \cdot p) \cdot (n \cdot p - n_i^2) \quad (\text{C.7})$$

Die Augerkoeffizienten c_n und c_p liegen dabei in der Größenordnung von $10^{-31} \text{ cm}^6/\text{s}$.

Da die genannten Prozesse gleichberechtigt sind, allerdings aufgrund der unterschiedlichen Abhängigkeit von der Beleuchtungsstärke aufweisen, soll eine effektive Lebensdauer wie folgt abgeschätzt werden: Die Dichten der Elektronen und Löcher unter Beleuchtung ergeben sich durch Addition der Injektionsdichte Δn zu den Dunkelwerten n_0 und p_0 . Des Weiteren wird angenommen, dass das für die SRH-Rekombination verantwortliche Defektniveau in der Bandlückenmitte befindet und die Rekombinationslebensdauern an diesem Defekt für beide Ladungstypen gleich sind. Es ergibt sich für p-dotiertes Material:

$$R = R_{BB} + R_{SRH} + R_{Auger} = b \cdot (np - n_i^2) + \frac{1}{\tau_{SRH,eff}} \frac{n \cdot p - n_i^2}{n + p + 2 \cdot n_i} + (c_n \cdot n + c_p \cdot p) \cdot (n \cdot p - n_i^2)$$

$$R = b \cdot \Delta n \cdot \left(\frac{p_0^2 + n_i^2}{p_0} + \Delta n \right) + \frac{\Delta n}{\tau_{SRH,eff}} \cdot \frac{p_0^2 + n_i^2 + p_0 \cdot \Delta n}{p_0^2 + n_i^2 + 2 \cdot p_0 \cdot (\Delta n + n_i)} + c \cdot \Delta n \cdot \left(\frac{p_0^2 + n_i^2}{p_0} + \Delta n \right) \cdot \left(\frac{p_0^2 + n_i^2}{p_0} + 2 \cdot \Delta n \right)$$

$$R \approx b \cdot \Delta n \cdot (p_0 + \Delta n) + \frac{\Delta n}{\tau_{SRH,eff}} \cdot \frac{p_0 + \Delta n}{p_0 + 2 \cdot (\Delta n + n_i)} + c \cdot \Delta n \cdot (p_0 + \Delta n) \cdot (p_0 + 2 \cdot \Delta n) \quad (\text{C.8})$$

Die Näherung ergibt sich hierbei unter der Annahme üblicher Werte für die Dotierung, dass heißt das p_0 viel grösser als n_i ist. Die effektive Lebensdauer, definiert als Quotient von Injektionsdichte und Rekombinationsrate, ist in Abbildung C.3 dargestellt. Offenbar wird das Rekombinationsverhalten fast ausschließlich durch SRH-Prozesse limitiert und stellt sich als eher gleichmäßig über einen weiten Injektionsdichtenbereich dar, was die Nutzung einer konstanten effektiven Lebensdauer für die weiteren Betrachtungen motiviert.

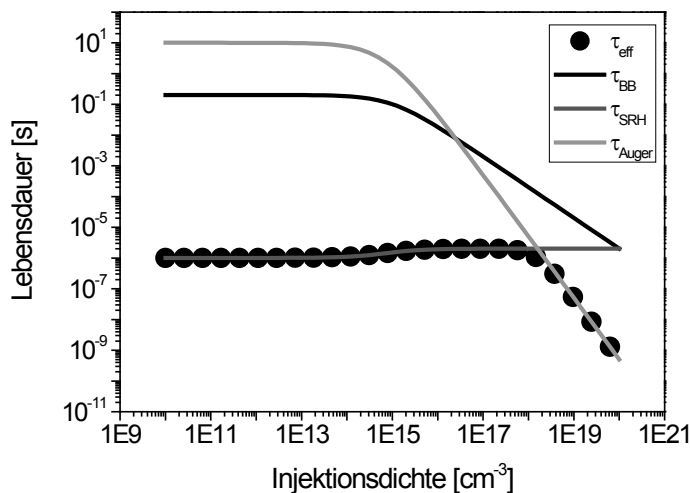


Abbildung C.3: Abhängigkeit der Lebensdauer von generierten Elektron-Loch-Paaren von der Injektionsdichte. Offensichtlich ist die effektive Lebensdauer unter den genannten Annahmen über einen weiten Bereich durch die SRH-Rekombination limitiert, Band-Band-Übergänge spielen eine untergeordnete Rolle und Auger-Prozesse zeigen erst bei sehr starker Beleuchtung einen Einfluss.

Um allerdings den Zusammenhang zwischen Photonenfluß und Injektionsdichte zu ergründen, ist es notwendig, Umverteilungsprozesse zu betrachten.

C.2 Mathematische Beschreibung des Ladungsträgerflusses

Ein grundlegendes Phänomen in der Natur ist die sogenannte Diffusion, ein Konzentrationsausgleich, welcher durch Gradienten getrieben wird. Allgemein werden diese Vorgänge häufig mit Hilfe einer Kontinuitätsgleich beschrieben:

$$\operatorname{div}(\vec{j}) + \frac{\partial \rho}{\partial t} = R \quad (\text{C.9})$$

Es handelt sich also um eine partielle Differentialgleichung, die die zeitliche Änderung einer Dichte ρ mit der Reaktionsrate R und dem ortsabhängigen Fluss j verknüpft. Anwendung findet diese Formulierung (für verschwindende Reaktionsrate) bei der Beschreibung von Erhaltungsgrößen wie Ladung oder Masse.

Da die Stromdichte gemäß des Fickschen Gesetzes als proportional zum Gradienten der Dichte angenommen werden kann, ergibt sich bei eindimensionaler Betrachtung:

$$-D \cdot \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial \rho}{\partial t} = R \quad (\text{C.10})$$

In obiger Gleichung stellt D die Diffusionskonstante dar. Eine Anpassung auf das Problem photo-generierter Elektron-Loch-Paare ergibt schließlich:

$$\frac{\partial \Delta n}{\partial t} = D_n \cdot \frac{\partial^2 \Delta n}{\partial x^2} + G_0 e^{-\alpha x} - \frac{\Delta n}{\tau_{eff}} \quad (\text{C.11})$$

Eine quasi Gleichgewichtslösung ergibt sich, wenn die rechte Seite der Gleichung verschwindet. Unter der Annahme einer unendlich dicken Probe und Einführung des Photonenflusses an der Oberfläche Φ_0 ergibt sich:

$$\Delta n(x) = \frac{\Phi_0 \cdot \tau_{eff}}{L} \cdot \frac{\alpha L}{\alpha^2 L^2 - 1} \cdot \left(\frac{s + \alpha D}{\frac{D}{L} + s} \cdot e^{-\frac{x}{L}} - e^{-\alpha x} \right) \quad (\text{C.12})$$

Die hierbei eingeführte Größe L ist die sogenannte Diffusionslänge, die sich aus der Wurzel des Produkts der Diffusionskonstante und der Lebensdauer ergibt. Der erste Term in Gleichung (C.12) erlaubt im Übrigen eine grobe Schätzung der Injektionsdichte für verschwindende Oberflächenrekombination und soll im Folgenden mit Δn_0 bezeichnet werden. Die Ortsabhängigkeit der Dichte der generierten Ladungsträger ist für Silizium unter Annahme der Beleuchtung mit Licht der Wellenlänge 650 nm und für verschiedene Oberflächenrekombinationsgeschwindigkeiten in Abbildung C.4 illustriert. Für verschwindende Oberflächenrekombination ergibt sich ein exponentieller Verlauf mit einer Abklingkonstante, die der Diffusionslänge entspricht. Für sehr große Aktivität der Oberfläche stellt sich ein deutliches Maximum der Injektionsdichte ein, begleitet von einem Absinken der durchschnittlichen Dichte. Der Einfluss der Aktivität auf die oberflächennahe Ladungsträgerdichte ist im rechten Graphen dargestellt.

Es bleibt hinzuweisen, dass eine Lösung der Diffusionsgleichung in höheren Dimensionen im Allgemeinen nicht analytisch möglich ist und man auf die Nutzung numerischer Hilfsmittel angewiesen ist. Als Beispiele sollen hier eine unstetige Beleuchtung, etwa Schattenwurf oder die Nutzung von fokussiertem Licht, und der Einfluss einer zweiten Oberfläche genutzt werden. Die Berechnungen erfolgen durch die Methode der finiten Elemente, angewendet im kommerziellen Paket „Comsol Multiphysics“.

Die Lösungen (Abbildung C.5) und extrahierten Verläufe entlang ausgezeichneter Richtungen (Abbildung C.6) zeigen deutlich das komplexe Verhalten des Systems: Während in der Tiefe (z-Achse) der Verlauf der Ladungsträgerdichten den bereits besprochenen eindimensionalen Lösungen folgt, stellt sich lateral ein komplizierter Verlauf ein. Im Inneren des beleuchteten Gebietes bleibt die Dichte fast konstant, außerhalb ist auch hier ein exponentieller Verlauf zu finden. Wichtig ist offenbar auch die Aktivität der Oberflächen in x-Richtung, wie die graue Spur verdeutlicht.

Der belichtete Halbleiter

Um allerdings den Grad der Beschreibung, vor allem bei Inhomogenität, zu verbessern, muss zusätzlich noch die Beschreibung der elektrostatischen Wechselwirkungen genutzt werden. Man erhält das sogenannte Drift-Diffusions-Modell.

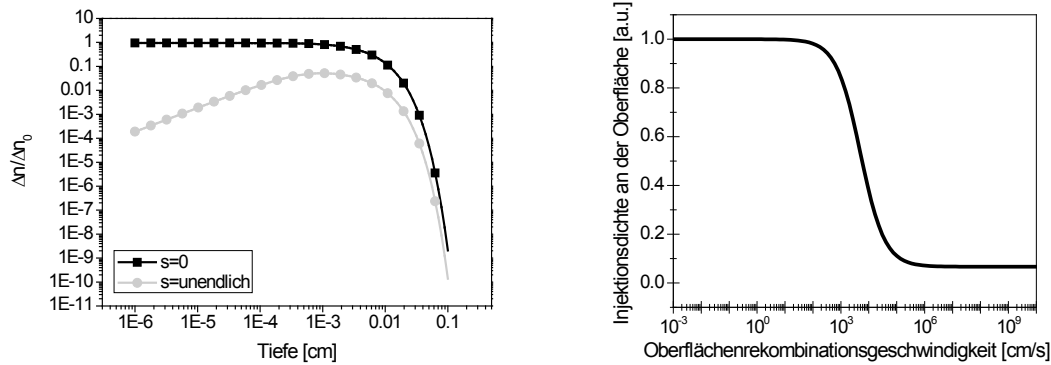


Abbildung C.4: Ortsabhängigkeit der normierten Injektionsdichte für verschiedenen Oberflächenaktivitäten (links). Der Einflussbereich hat hier eine Länge von etwa 100 µm, danach fällt die Dichte stark ab. Die Auswirkungen unterschiedlicher Oberflächenrekombination sind offensichtlich und bewirken eine starke Änderung der Dichte vor allem nahe der Oberfläche (rechts).

C.3 Das Drift-Diffusions-Modell

Hierfür wird zusätzlich zur Diffusionskomponente des Stroms eine Driftkomponente eingefügt, welche als proportional zum elektrischen Feld angenommen wird:

$$j_{\text{Drift}} = q\rho\mu E \quad (\text{C.13})$$

In Gleichung (C.13) steht q für die Elementarladung, ρ für die jeweilige Ladungsträgerdichte, μ für die Beweglichkeit und E schließlich für das elektrische Feld. Zusammen mit der Poissongleichung (siehe Anhang A) müssen also drei gekoppelte Differentialgleichungen gelöst werden:

$$\begin{aligned} \text{div} E &= \frac{q}{\varepsilon} (p - n + N_D - N_A) \\ \text{div}(qn\mu_n E + qD_n \nabla n) &= R \\ \text{div}(-qp\mu_p E + qD_p \nabla p) &= R \end{aligned} \quad (\text{C.14})$$

In obiger Gleichung treten zusätzlich zu den die Diffusion beschreibenden Größen noch die Dotierungsdichten N_A und N_D auf, die Rekombinationsrate R entnimmt man den bereits erklärten Modellen. Die Lösung dieses Systems ist nicht analytisch möglich, vor allem nicht, wenn experimentelle Bedingungen beschrieben werden sollen. Beispielhaft soll in dieser Stelle ein wichtiger Fall besprochen werden: die Ausbildung einer Photospannung an der Oberfläche und deren Abhängigkeiten.

In Frage kommende Parameter sind dabei: Die Beleuchtungsstärke, die Lebensdauern an Oberfläche und im Volumen und die Dotierungsdichte. Vereinfachend wird Dotierung nur mit einer Sorte Dotanden (hier p-dotiert) angenommen, sowie die Lebensdauern für Löcher und Elektronen gleich angesetzt.

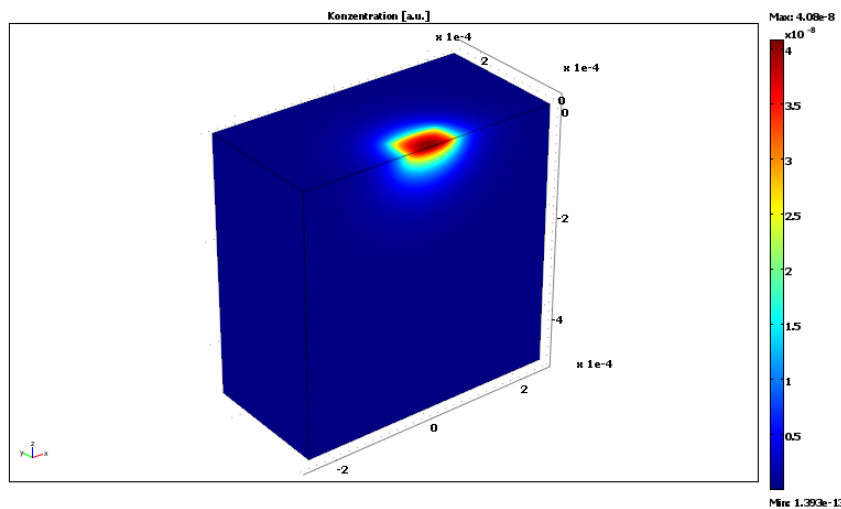


Abbildung C.5: Ladungsträgerdichte erzeugt durch ungleichmäßige Beleuchtung. Die Lichtquelle bestrahlt einen quadratischen Bereich mit 100 μm Kantenlänge.

Für ungleichmäßige Beleuchtung (hier nur im Bereich $x < 0$) ergibt sich das in Abbildung C.7 dargestellte typische Bild: In den bestrahlten Bereichen steigt das Potential gegenüber den unbestrahlten Bereichen an. Ursache dafür ist ein Ungleichgewicht der Löcher- und Elektronendichten, verursacht durch unterschiedliche Diffusionskonstanten. Der Effekt ist bekannt und wird im Allgemeinen unter dem Begriff Oberflächenphotospannung (SPV) geführt [51]. Der Unterschied der Potentiale zwischen hell und dunkel ist ebenfalls in Abbildung C.7 illustriert, er folgt folgender Form:

$$V(z=0) = V_0 \ln \left(1 + \frac{G}{G_T} \right) \quad (\text{C.15})$$

Der belichtete Halbleiter

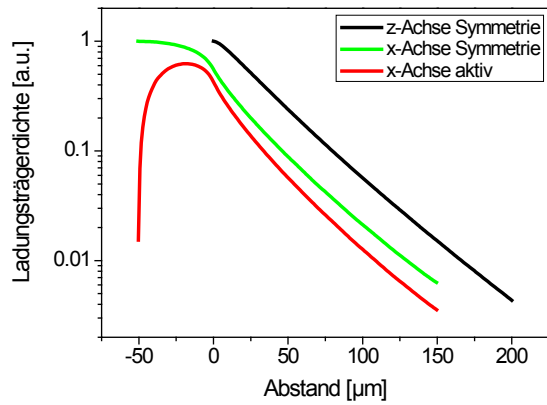


Abbildung C.6: Verlauf der normierten Ladungsträgerdichten entlang ausgezeichneter Richtungen. Offenbar gibt es sowohl in der Tiefe (z-Achse) als auch lateral (x-Achse) einen großen Einfluss der entsprechenden Aktivität der Oberflächen.

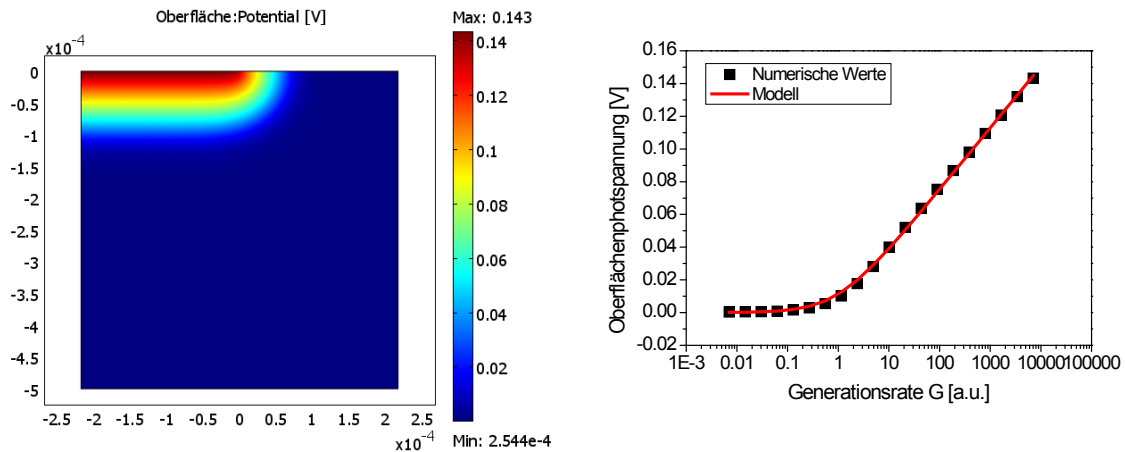


Abbildung C.7: Verteilung des Potentials in einer Siliziumprobe bei ungleichmäßiger Beleuchtung (links) und Verlauf des Potentials an der Oberfläche in Abhängigkeit der Beleuchtungsstärke (rechts). Offenbar existiert ein logarithmischer Zusammenhang gemäß (C.15).

Die in (C.15) auftretende Konstante V_0 entspricht dabei etwa der thermischen Spannung (25 mV bei Raumtemperatur) und G_T ist eine Schwellengenerationsrate. Ergebnisse parametrischer Studien zeigt Abbildung C.8.

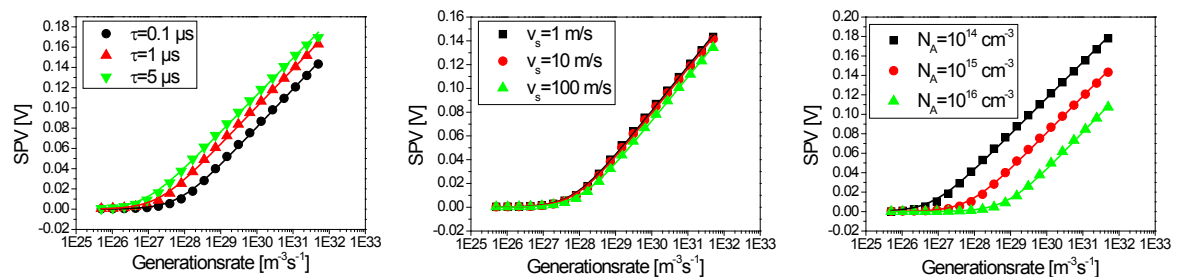


Abbildung C.8: Untersuchungen zum Einfluss verschiedener Materialparameter auf die Photospannung. Prinzipiell verschieben alle Werte, hier Lebensdauer, Rekombinationsgeschwindigkeit und Dotierung, den Schwellwert in Gleichung (C.15).

Motiviert durch Gleichung (A.5) lassen sich die gezeigten Ergebnisse wie folgt erklären: Da die Ladungsträgerdichte exponentiell mit der Spannung steigt, kann man im Umkehrschluss näherungsweise fordern:

$$V = V_i \ln \left(1 + \frac{\Delta n}{n} \right) \quad (\text{C.16})$$

Nutzt man weiterhin (C.12) unter Beachtung der üblichen Größenordnungen und entwickelt für kleine Rekombinationsgeschwindigkeiten s , ergibt sich:

$$V \approx V_i \ln \left(1 + \frac{\Phi \tau}{\Lambda N} \right); \frac{1}{\Lambda} = \frac{1}{L} - \frac{s}{D} \quad (\text{C.17})$$

Dafür wurde vereinfachend angenommen, dass die Ladungsträgerdichte ohne Beleuchtung etwa der Dotierungsdichte N entspricht. Die eingeführte Größe Λ hat die Einheit einer Länge und stellt eine effektive Diffusionslänge dar. (C.17) entspricht der Darstellung in (C.15), wobei ein Schwellfluss Φ_T wie folgt definiert werden kann:

$$\Phi_T = \frac{\Lambda N}{\tau} \quad (\text{C.18})$$

Zum Abschluss sollen noch typische Werte für die Photospannung geschätzt werden. Dazu werden die tatsächlich im genutzten System auftretenden Parameter genutzt: Leistung des Detektionslasers im Bereich einiger mW, Wellenlänge des Lasers 650 nm, bestrahlte Fläche etwa $50 \mu\text{m} \times 10 \mu\text{m}$, Dotierungsdichten zwischen 10^{15} cm^{-3} und 10^{20} cm^{-3} , Lebensdauern im Bereich von μs . Die Oberflächenrekombinationsrate wird auf null gesetzt, was auch experimentell durch gezielte Oberflächenpassivierung möglich ist. Man erhält Photonenflüsse in der Größenordnung von 10^{25} pro Quadratmeter und Sekunde. Damit, natürlich abhängig von der Dotierung, Spannungen im Bereich zwischen 2 mV und 200 mV. Nur durch die Beleuchtung ist es also ohne weiteres möglich, das Kontaktpotential zwischen unterschiedlich dotierten Bereichen auszugleichen.

D Multipolentwicklungen

Eine genauere Modellierung, verglichen mit einfachen Punktladungsmodellen, des von der Spitze generierten Feldes ist möglich, wenn zusätzlich zu der für große Abstände gültigen Punktladungsnäherung höhere elektrostatische Elemente benutzt werden. Allgemein gilt dabei:

$$\Phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{Q}{r} + \frac{\mathbf{r} \cdot \mathbf{p}}{r^3} + \frac{1}{2} \sum_{k,l} Q_{kl} \frac{r_k r_l}{r^5} + \dots \right) \quad (\text{D.1})$$

In Gleichung (D.1) steht $\Phi(\mathbf{r})$ für das elektrostatische Potential im Punkt \mathbf{r} , Q entspricht einer Punktladung, \mathbf{p} einem Dipolmoment und die Koeffizienten Q_{kl} sind Quadrupolmomente. Durch Ableitung ergibt sich das elektrische Feld (bei Abschneiden der Reihe nach dem Dipolbeitrag) zu:

$$\mathbf{E}(\mathbf{r}) = -\nabla\Phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{Q}{r^3} \mathbf{r} + \frac{3\mathbf{p} \cdot \mathbf{r}}{r^5} \mathbf{r} - \frac{\mathbf{p}}{r^3} \right) \quad (\text{D.2})$$

Beachtet man weiterhin, dass im Falle eines geladenen Körpers vor einer geerdeten Fläche die Methode der Bildladungen benutzt werden kann um die Randbedingungen zu erfüllen, erhält man:

$$\mathbf{E}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \left(\frac{Q_+}{r_+^3} \mathbf{r}_+ + \frac{3\mathbf{p}_+ \cdot \mathbf{r}_+}{r_+^5} \mathbf{r}_+ - \frac{\mathbf{p}_+}{r_+^3} \right) + \frac{1}{4\pi\epsilon_0} \left(\frac{Q_-}{r_-^3} \mathbf{r}_- + \frac{3\mathbf{p}_- \cdot \mathbf{r}_-}{r_-^5} \mathbf{r}_- - \frac{\mathbf{p}_-}{r_-^3} \right) \quad (\text{D.3})$$

Die Bedeutung der mit „+/-“, indizierten Größen entnimmt man Abbildung D.1.

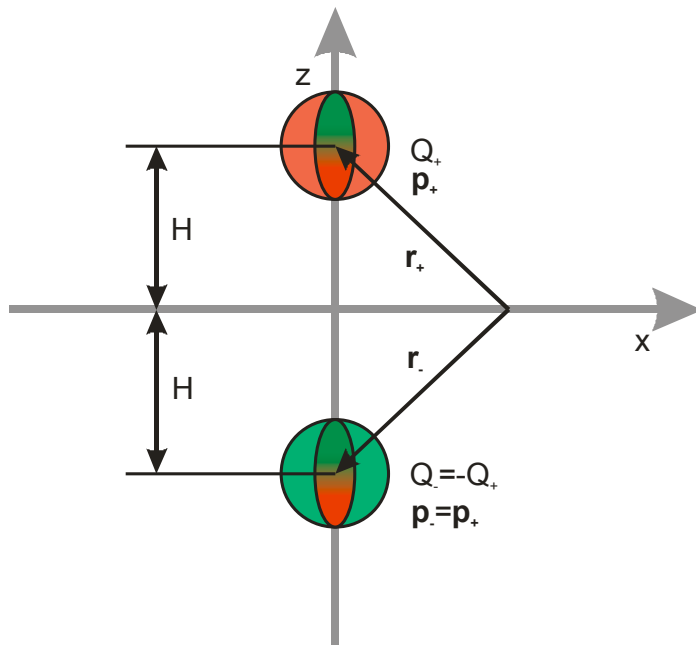


Abbildung D.1: Geometrie der Multipolentwicklung. Die freien Parameter \mathbf{p} und Q müssen so gewählt werden, dass die Äquipotentialflächen Kugeln mit dem geforderten Radius sind.

Für einen Punkt auf der x-Achse gilt somit:

Multipolentwicklungen

$$E_z(x) = \frac{1}{4\pi\epsilon_0} \left(\frac{2Q_+H}{(x^2 + H^2)^{\frac{3}{2}}} + \frac{6pH^2}{(x^2 + H^2)^{\frac{5}{2}}} - \frac{2p}{(x^2 + H^2)^{\frac{3}{2}}} \right) \quad (D.4)$$

Beachtet man weiterhin, dass die Ladung einer isolierten Kugel mit Radius R mit Potential V sich leicht zu $Q=4\pi\epsilon_0RV$ ermitteln lässt, ergibt sich:

$$E_z(x) = \frac{1}{4\pi\epsilon_0} \left(\frac{2H \cdot 4\pi\epsilon_0RV}{(x^2 + H^2)^{\frac{3}{2}}} + \frac{6H^2 \cdot 4\pi\epsilon_0R^2V_p}{(x^2 + H^2)^{\frac{5}{2}}} - \frac{2 \cdot 4\pi\epsilon_0R^2V_p}{(x^2 + H^2)^{\frac{3}{2}}} \right) \quad (D.5)$$

$$E_z(x) = \frac{2HRV}{(x^2 + H^2)^{\frac{3}{2}}} + \frac{6H^2R^2V_p}{(x^2 + H^2)^{\frac{5}{2}}} - \frac{2R^2V_p}{(x^2 + H^2)^{\frac{3}{2}}} = \frac{2HRV - 2R^2V_p}{(x^2 + H^2)^{\frac{3}{2}}} + \frac{6H^2R^2V_p}{(x^2 + H^2)^{\frac{5}{2}}}$$

In Gleichung (D.5) wurde das Dipolmoment p durch $4\pi\epsilon_0R^2V_p$ ersetzt. Eine Normierung der Koordinaten auf R und der Potentiale auf V liefert schließlich:

$$E_z(x) = \frac{2\hat{H}V - 2\hat{V}_pV}{R(\hat{x}^2 + \hat{H}^2)^{\frac{3}{2}}} + \frac{6\hat{H}^2\hat{V}_pV}{R(\hat{x}^2 + \hat{H}^2)^{\frac{5}{2}}} = 2\frac{V}{R} \left(\frac{\hat{H} - \hat{V}_p}{(\hat{x}^2 + \hat{H}^2)^{\frac{3}{2}}} + \frac{3\hat{H}^2\hat{V}_p}{(\hat{x}^2 + \hat{H}^2)^{\frac{5}{2}}} \right) \quad (D.6)$$

Um einen Ausdruck für die Abhängigkeit der auf ein Flächenelement wirkenden Kraft zu erhalten, muss das elektrische Feld quadriert werden (der Einfachheit halber werden die Konstanten zusammengefasst zu C):

$$F(x) = C \left(\frac{\hat{H} - \hat{V}_p}{(\hat{x}^2 + \hat{H}^2)^{\frac{3}{2}}} + \frac{3\hat{H}^2\hat{V}_p}{(\hat{x}^2 + \hat{H}^2)^{\frac{5}{2}}} \right)^2 = C \left(\frac{(\hat{H} - \hat{V}_p)^2}{(\hat{x}^2 + \hat{H}^2)^3} + \frac{6(\hat{H} - \hat{V}_p)\hat{H}^2\hat{V}_p}{(\hat{x}^2 + \hat{H}^2)^4} + \frac{9\hat{H}^4\hat{V}_p^2}{(\hat{x}^2 + \hat{H}^2)^5} \right) \quad (D.7)$$

Im Rahmen der Multipolnäherung erhält man somit zusätzlich zu der aus grundlegenden Überlegungen erkennbaren Abhängigkeit (erster Term) Terme mit höheren Exponenten.

Um eine Abschätzung des Einflussbereichs, ein Äquivalent zur Auflösung in den Experimenten, zu erhalten, muss allerdings der Wert des Dipolmoments p beziehungsweise dessen Parameterabhängigkeit ermittelt werden. Dazu kann man von Gleichung (D.1) ausgehend das Potential an signifikanten Punkten festlegen und die erhaltenen Gleichungen lösen. Diese Prozedur liefert, angewendet auf $(0,0,H-R)$ und $(0,0,H+R)$:

$$p = -4\pi\epsilon_0V \frac{\left(\frac{1}{2H-R} - \frac{1}{2H+R} \right)}{\left(\frac{1}{R} - \frac{1}{2H-R} \right) \left(\frac{1}{R^2} + \frac{1}{(2H+R)^2} \right) + \left(\frac{1}{R} - \frac{1}{2H+R} \right) \left(\frac{1}{R^2} - \frac{1}{(2H-R)^2} \right)} \quad (D.8)$$

Das normierte Dipolpotential ist damit:

$$\hat{V}_p = - \frac{\left(\frac{1}{2\hat{H}-1} - \frac{1}{2\hat{H}+1} \right)}{\left(1 - \frac{1}{2\hat{H}-1} \right) \left(1 + \frac{1}{(2\hat{H}+1)^2} \right) + \left(1 - \frac{1}{2\hat{H}+1} \right) \left(1 - \frac{1}{(2\hat{H}-1)^2} \right)} \quad (\text{D.9})$$

Das Ergebnis für das normierte Potential der Punktladung, \hat{V}_Q , lautet:

$$\hat{V}_Q = \frac{\left(2 + \frac{1}{(2\hat{H}+1)^2} - \frac{1}{(2\hat{H}-1)^2} \right)}{\left(1 - \frac{1}{2\hat{H}-1} \right) \left(1 + \frac{1}{(2\hat{H}+1)^2} \right) + \left(1 - \frac{1}{2\hat{H}+1} \right) \left(1 - \frac{1}{(2\hat{H}-1)^2} \right)} \quad (\text{D.10})$$

Eine Darstellung der gewonnenen Ausdrücke als Funktion der normierten Höhe ist in Abbildung D.2 gegeben. Deutlich zu erkennen ist die starke Abhängigkeit des Dipolpotentials ($\sim H^{-8}$ für kleine H und $\sim H^{-2}$ für große H). Im Gegensatz dazu konvergiert das Potential der Punktladung sehr schnell gegen den Wert, der für eine isolierte Kugel erwartet wird (hier 1).

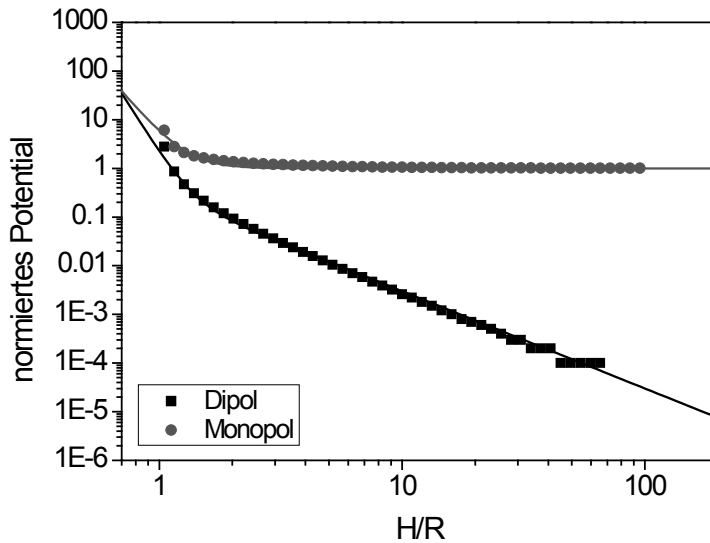


Abbildung D.2: Höhenabhängigkeit der normierten Potentiale. Offenbar muss der Einfluss des Dipolmoments nur für Höhen kleiner dem doppelten Radius der Kugel beachtet werden. Die Symbole entsprechen dabei den Ergebnissen der Rechnung, die Linien entsprechen Potenzgesetzen.

Werden die normierten Potentiale in Gleichung (D.7) eingesetzt und quadratische Terme in \hat{V}_p unterdrückt, ergibt sich die in Abbildung D.3 illustrierte Abhängigkeit der Maximalkraft, wirksam am Ort $x=0$, von der Höhe. Offensichtlich ist der Einfluss des Dipols wiederum nur für kleine Abstände relevant, was sich durch einen Abfall der Kraft proportional zu hohen Potenzen von H bemerkbar macht.

Zusammenfassend kann somit formuliert werden, dass es im Fall makroskopischer Objekte sinnvoll erscheint, auftretende Kräfte über dem Abstand zwischen Objekt und der Position von

Multipolentwicklungen

Interesse zu parametrisieren. Es ist zu erwarten, dass dieser Abstand in kleinen Potenzen in die Rechnung eingeht, wie Gleichung (D.7) zeigt.

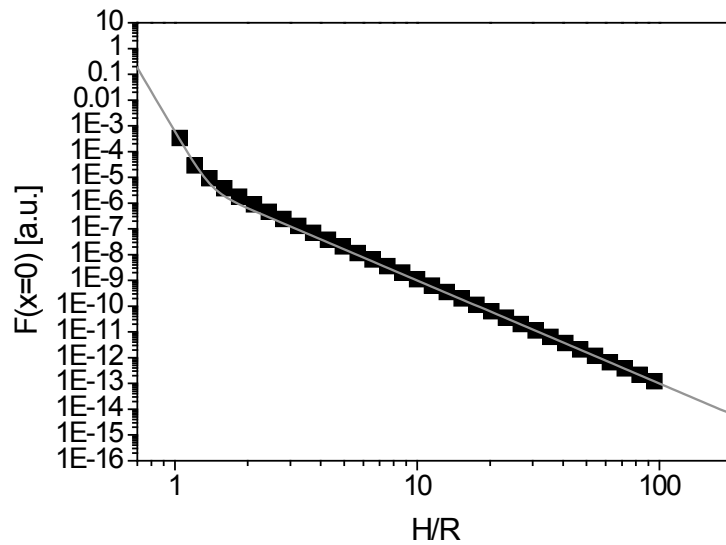


Abbildung D.3: Abhängigkeit der Maximalkraft (Symbole), wirksam am Punkt $x=0$, im Rahmen der Dipolnäherung für eine Kugel. Für Höhen im Bereich des Radius spielt der Dipolanteil eine Rolle, für größere Werte überwiegt der Monopolanteil. Die ebenfalls eingezeichnete Linie ist eine Summe von Potenzgesetzen mit unterschiedlichen Exponenten, hier 4 und 16.

Literaturliste

1. Abbe, E. and O. Lummer, *Die Lehre von der Bildentstehung im Mikroskop* 1910, Braunschweig: Vieweg. XII, 108 S.
2. Knoll, M. and E. Ruska, *Das Elektronenmikroskop*. Zeitschrift für Physik A Hadrons and Nuclei, 1932. **78**(5): p. 318-339.
3. Binnig, G., et al., *Tunneling through a Controllable Vacuum Gap*. Applied Physics Letters, 1982. **40**(2): p. 178-180.
4. Binnig, G., C.F. Quate, and C. Gerber, *Atomic Force Microscope*. Physical Review Letters, 1986. **56**(9): p. 930-933.
5. ITRS. *The International Technology Roadmap for Semiconductors*. 2010; Available from: <http://www.itrs.net/home.html>.
6. Wikipedia. *Rastertunnelmikroskop*. Available from: <http://de.wikipedia.org/wiki/Rastertunnelmikroskop>.
7. Timoshenko, S.P., *History of strength of materials : with a brief account of the history of theory of elasticity and theory of structures* 1983, New York: Dover.
8. Martin, Y., C.C. Williams, and H.K. Wickramasinghe, *Atomic Force Microscope Force Mapping and Profiling on a Sub 100-A Scale*. Journal of Applied Physics, 1987. **61**(10): p. 4723-4729.
9. Martin, Y., D.W. Abraham, and H.K. Wickramasinghe, *High-Resolution Capacitance Measurement and Potentiometry by Force Microscopy*. Applied Physics Letters, 1988. **52**(13): p. 1103-1105.
10. Zhong, Q., et al., *Fractured Polymer Silica Fiber Surface Studied by Tapping Mode Atomic-Force Microscopy*. Surface Science, 1993. **290**(1-2): p. L688-L692.
11. Nonnenmacher, M., M.P. Oboyle, and H.K. Wickramasinghe, *Kelvin Probe Force Microscopy*. Applied Physics Letters, 1991. **58**(25): p. 2921-2923.
12. Sauvigny, F., *Foundations and integral representations*. Partial differential equations ; 2006, Berlin: Springer. XIV, 437 S.
13. Zerweck, U., et al., *Accuracy and resolution limits of Kelvin probe force microscopy*. Physical Review B, 2005. **71**(12).
14. Mathworks. *Matlab Documentation*. 2011; Available from: <http://www.mathworks.de/help/techdoc/>.
15. Press, W.H., *Numerical Recipes 3rd Edition: The Art of Scientific Computing* 2007: Cambridge University Press.
16. Davis, T.A., *Algorithm 832: UMFPACK V4.3---an unsymmetric-pattern multifrontal method*. ACM Trans. Math. Softw., 2004. **30**(2): p. 196-199.
17. Hestenes, M.R. and E. Stiefel, *Methods of Conjugate Gradients for Solving Linear Systems*. Journal of Research of the National Bureau of Standards, 1952. **49**(6): p. 409-436.
18. Steinke, P., *Finite-Elemente-Methode : rechnergestützte Einführung ; mit 39 Tabellen*. 2., neu bearb. Aufl.. ed 2007, Berlin: Springer. XV, 391 S.
19. Jackson, J.D., *Classical electrodynamics*. 3. ed 1999, New York: Wiley. XXI, 808 S.
20. Jacobs, H.O., et al., *Resolution and contrast in Kelvin probe force microscopy*. Journal of Applied Physics, 1998. **84**(3): p. 1168-1173.
21. Ratzke, M. and H. Reif, *On the reliability of scanning probe based electrostatic force measurements*. Microelectronic Engineering, 2007. **84**(3): p. 512-516.
22. Sze, S.M. and K.K. Ng, *Physics of semiconductor devices*. 3. ed 2007, Hoboken, NJ: Wiley. X, 815 S.
23. Nicollian, E.H. and J.R. Brews, *MOS (Metal oxide semiconductor) physics and technology* 1982, New York: Wiley. XV, 906 S.

Literaturliste

24. Williams, C.C., et al., *Lateral Dopant Profiling on a 100 nm Scale by Scanning Capacitance Microscopy*. Journal of Vacuum Science & Technology a-Vacuum Surfaces and Films, 1990. **8**(2): p. 895-898.
25. Raineri, V. and F. Giannazzo, *Scanning capacitance microscopy of semiconductor materials*. Beam Injection Assessment of Microstructures in Semiconductors 2000, 2000. **78-79**: p. 425-432.
26. Hudlet, S., et al., *Evaluation of the capacitive force between an atomic force microscopy tip and a metallic surface*. The European Physical Journal B - Condensed Matter and Complex Systems, 1998. **2**(1): p. 5-10.
27. Durand, É., *Electrostatique*, Paris: Masson.
28. Evans, L.C., *Partial differential equations*. Reprinted with corr.. ed. Graduate studies in mathematics, 2008, Providence, RI: American Math. Soc. XVII, 662 S.
29. Stein, E.M. and R. Shakarchi, *Complex analysis*. Princeton lectures in analysis ;2003, Princeton, N.J.: Princeton Univ. Press. XVII, 379 S.
30. Shannon, C.E., *Communication in the Presence of Noise*. Proceedings of the Institute of Radio Engineers, 1949. **37**(1): p. 10-21.
31. Ratzke, M., B. Yang, and J. Reif, *Fixed charge density distribution on pulsed laser deposited high-k oxide layers*. physica status solidi (c), 2008. **5**(10): p. 3304-3308.
32. Muller, K., et al., *Scanning Kelvin probe and photoemission electron microscopy of organic source-drain structures*. Synthetic Metals, 2004. **146**(3): p. 377-382.
33. Ratzke, M., et al., *Electrical properties of laser-ablation-initiated self-organized nanostructures on silicon surface*. Materials Science and Engineering B-Solid State Materials for Advanced Technology, 2006. **134**(2-3): p. 114-117.
34. Ratzke, M., et al., *Scanning probe studies of amorphous silicon subjected to laser annealing*. physica status solidi (c), 2011. **8**(4): p. 1351-1355.
35. Magonov, S.N., V. Elings, and M.H. Whangbo, *Phase imaging and stiffness in tapping-mode atomic force microscopy*. Surface Science, 1997. **375**(2-3): p. L385-L391.
36. Kittler, M., et al., *Self-organized pattern formation of biomolecules at silicon surfaces: Intended application of a dislocation network*. Materials Science & Engineering C-Biomimetic and Supramolecular Systems, 2006. **26**(5-7): p. 902-910.
37. Kittler, M., et al., *Properties of dislocation networks formed by Si wafer direct bonding*. Materials Science in Semiconductor Processing, 2006. **9**(1-3): p. 96-101.
38. Wilshaw, P.R., Booker, G.R., *New Results and an Interpretation for SEM EBIC Contrast arising from individual Dislocations in Silicon*. Institute of Physics Conference Series, 1985(76): p. 329-336.
39. Shikler, R., et al., *Measuring minority-carrier diffusion length using a Kelvin probe force microscope*. Physical Review B, 2000. **61**(16): p. 11041-11046.
40. Tyagi, M.S. and R. Vanoverstraeten, *Minority-Carrier Recombination in Heavily-Doped Silicon*. Solid-State Electronics, 1983. **26**(6): p. 577-597.
41. Kim, D.M., et al., *Deep-depletion high-frequency capacitance-voltage responses under photonic excitation and distribution of interface states in MOS capacitors*. IEEE Transactions on Electron Devices, 2003. **50**(4): p. 1131-1134.
42. Buh, G.H., et al., *Factors influencing the capacitance-voltage characteristics measured by the scanning capacitance microscope*. Journal of Applied Physics, 2003. **94**(4): p. 2680-2685.
43. Zaumseil, P., M. Birkholz, and G. Weidner, *X-ray characterization of periodic sub-nm surface relief gratings*. physica status solidi (a), 2007. **204**(8): p. 2657-2661.
44. Ziegler, J.F., *SRIM*, 2011: <http://www.srim.org>.
45. London, F., *The general theory of molecular forces*. Transactions of the Faraday Society, 1937. **33**: p. 8b-26.

46. Landau, L.D. and E.M. Lifshitz, *Theory of elasticity*. 3. ed. Course of theoretical physics ;1986, London: Pergamon Press. 134 S.
47. Derjaguin, B.V., V.M. Muller, and Y.P. Toporov, *Effect of Contact Deformations on Adhesion of Particles*. Journal of Colloid and Interface Science, 1975. **53**(2): p. 314-326.
48. Jellison, G.E. and F.A. Modine, *Optical-Absorption of Silicon between 1.6-eV and 4.7-eV at Elevated-Temperatures*. Applied Physics Letters, 1982. **41**(2): p. 180-182.
49. Chelikowsky, J.R. and M.L. Cohen, *Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors*. Physical Review B, 1976. **14**(2): p. 556.
50. Arguirov, T.V., *Electro-optical properties of dislocations in silicon and their possible application for light emitters*, 2007, BTU Cottbus: Cottbus.
51. Schroder, D.K., *Surface voltage and surface photovoltage: history, theory and applications*. Measurement Science & Technology, 2001. **12**(3): p. R16-R31.